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

The data has some good news. China and South Korea have managed to remarkably bring down the growth in 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 experience 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 coutries seem to be following more the Italian model rather than South Korea.

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 exposire, the new data seem to good to be true.

In summary, there is an inkling of good news in the Italian data. Let’s see if the trend continues. South Korea and China (if the data are correct) shows that the virus can be beat, BUT 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. Growth rates were determined as:

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

# Recommended Actions

*Disclaimer*: I am not an epidemiologist so my analysis is very much one of an armchair scientist. However, I am an actuary, so I know a thing or two about risk rates.

The data is also very messy and is dependent on differing testing regimes rather than true infection experience.

# Data Sources

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

# Initialize Data

## Warning: package 'rstan' was built under R version 3.6.3

## Warning: package 'StanHeaders' was built under R version 3.6.2

## Warning: package 'ggplot2' was built under R version 3.6.3

## Warning: package 'dplyr' was built under R version 3.6.3

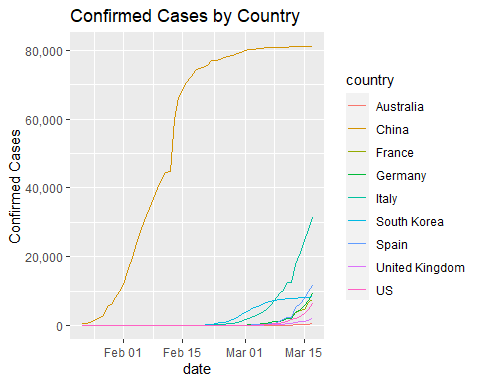
## Warning: package 'tidyr' was built under R version 3.6.3

## Warning: package 'bayesplot' was built under R version 3.6.3

## Warning: package 'corrplot' was built under R version 3.6.3

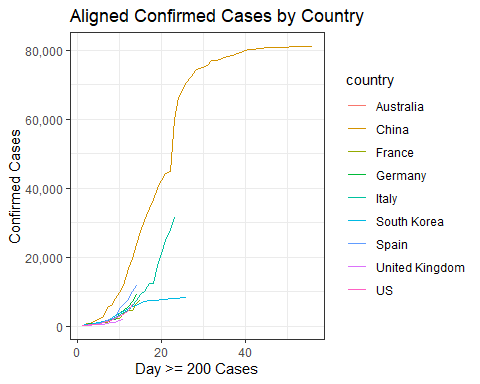
## Warning: Removed 1 rows containing missing values.  
  
## Warning: Removed 1 rows containing missing values.

# Data Exploration



Correlation Analysis

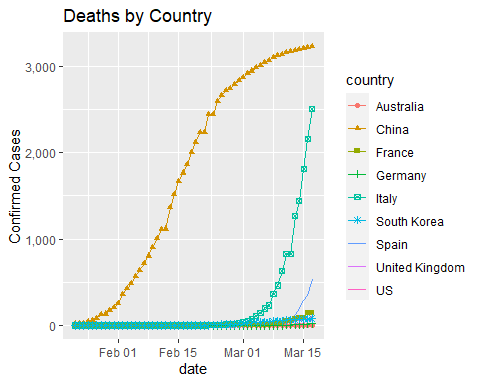
## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.



# Analysis of Deaths

## Warning: The shape palette can deal with a maximum of 6 discrete values because  
## more than 6 becomes difficult to discriminate; you have 9. Consider  
## specifying shapes manually if you must have them.

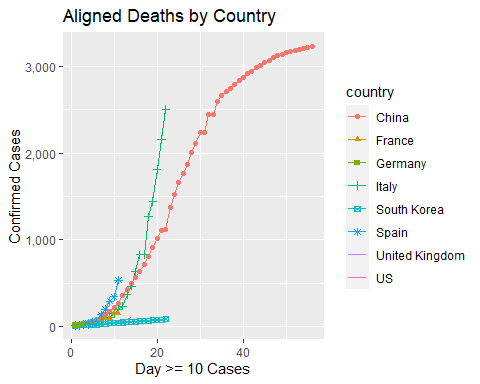
## Warning: Removed 168 rows containing missing values (geom\_point).



## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.

## Warning: The shape palette can deal with a maximum of 6 discrete values because  
## more than 6 becomes difficult to discriminate; you have 8. Consider  
## specifying shapes manually if you must have them.

## Warning: Removed 18 rows containing missing values (geom\_point).



# Gaussian Process Analysis

## Prepare input

*Confirmed Cases*

Deaths

# Stan Model with Block Diagonal of Confirmed Cases

## Warning: There were 2 divergent transitions after warmup. Increasing adapt\_delta above 0.8 may help. See  
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup

## Warning: Examine the pairs() plot to diagnose sampling problems

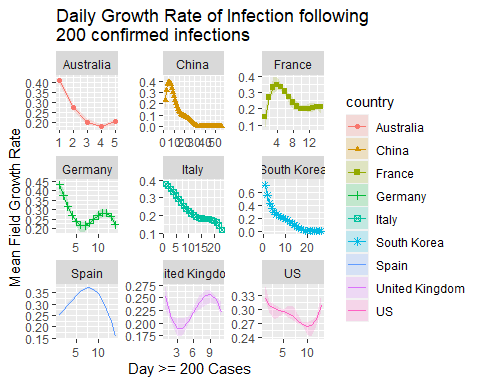
## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be unreliable.  
## Running the chains for more iterations may help. See  
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quantiles may be unreliable.  
## Running the chains for more iterations may help. See  
## http://mc-stan.org/misc/warnings.html#tail-ess

## mean se\_mean sd 2.5% 25%  
## eta 0.27036294 0.0041494068 0.062465362 0.18162750 0.22747090  
## rho[1] 16.35916734 0.2647534719 4.751551476 8.78317922 13.11427688  
## sigma[1] 0.15973510 0.0032288332 0.094800650 0.06434213 0.09977515  
## sigma[2] 0.07851696 0.0002195332 0.008063964 0.06500117 0.07282113  
## sigma[3] 0.19086307 0.0011268880 0.040988812 0.12901014 0.16132041  
## sigma[4] 0.16812699 0.0011193361 0.040234071 0.11058273 0.14002354  
## sigma[5] 0.07910169 0.0003387172 0.013840498 0.05743248 0.06949625  
## sigma[6] 0.09748065 0.0004607750 0.016950211 0.07070443 0.08588645  
## sigma[7] 0.22143410 0.0017057095 0.052279957 0.14710331 0.18477811  
## sigma[8] 0.18567208 0.0012747384 0.051395244 0.11542325 0.15092737  
## sigma[9] 0.11733310 0.0008042995 0.029508841 0.07496262 0.09687069  
## 50% 75% 97.5% n\_eff Rhat  
## eta 0.25999304 0.30339738 0.41325705 226.6243 1.0303003  
## rho[1] 15.79758274 18.86290279 27.88742101 322.0976 1.0153536  
## sigma[1] 0.13589004 0.18406350 0.41412950 862.0479 1.0067505  
## sigma[2] 0.07780975 0.08336770 0.09611407 1349.2630 1.0000972  
## sigma[3] 0.18451352 0.21343498 0.28812339 1323.0285 1.0001790  
## sigma[4] 0.16168900 0.18815491 0.26497304 1292.0131 1.0004168  
## sigma[5] 0.07694698 0.08634632 0.11069397 1669.6630 0.9980214  
## sigma[6] 0.09539667 0.10684186 0.13616700 1353.2322 1.0003761  
## sigma[7] 0.21217495 0.24879718 0.34068414 939.4211 1.0089904  
## sigma[8] 0.17542254 0.20955927 0.32104263 1625.5627 0.9992226  
## sigma[9] 0.11290281 0.13158463 0.19126575 1346.0732 1.0017847

## Warning: The shape palette can deal with a maximum of 6 discrete values because  
## more than 6 becomes difficult to discriminate; you have 9. Consider  
## specifying shapes manually if you must have them.

## Warning: Removed 38 rows containing missing values (geom\_point).



# Stan Model with Block Diagonal of Deaths

## Warning: There were 1 divergent transitions after warmup. Increasing adapt\_delta above 0.8 may help. See  
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup

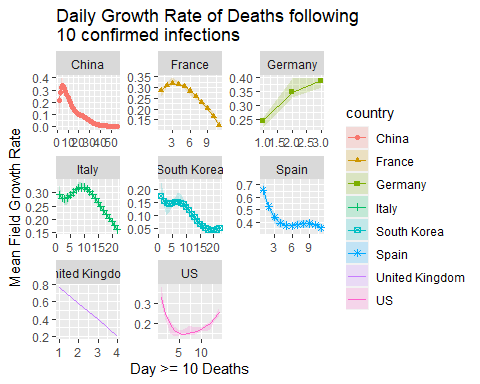
## Warning: Examine the pairs() plot to diagnose sampling problems

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be unreliable.  
## Running the chains for more iterations may help. See  
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quantiles may be unreliable.  
## Running the chains for more iterations may help. See  
## http://mc-stan.org/misc/warnings.html#tail-ess

## Warning: The shape palette can deal with a maximum of 6 discrete values because  
## more than 6 becomes difficult to discriminate; you have 8. Consider  
## specifying shapes manually if you must have them.

## Warning: Removed 18 rows containing missing values (geom\_point).



# Run LKJ Stan Model

Correlation Analysis