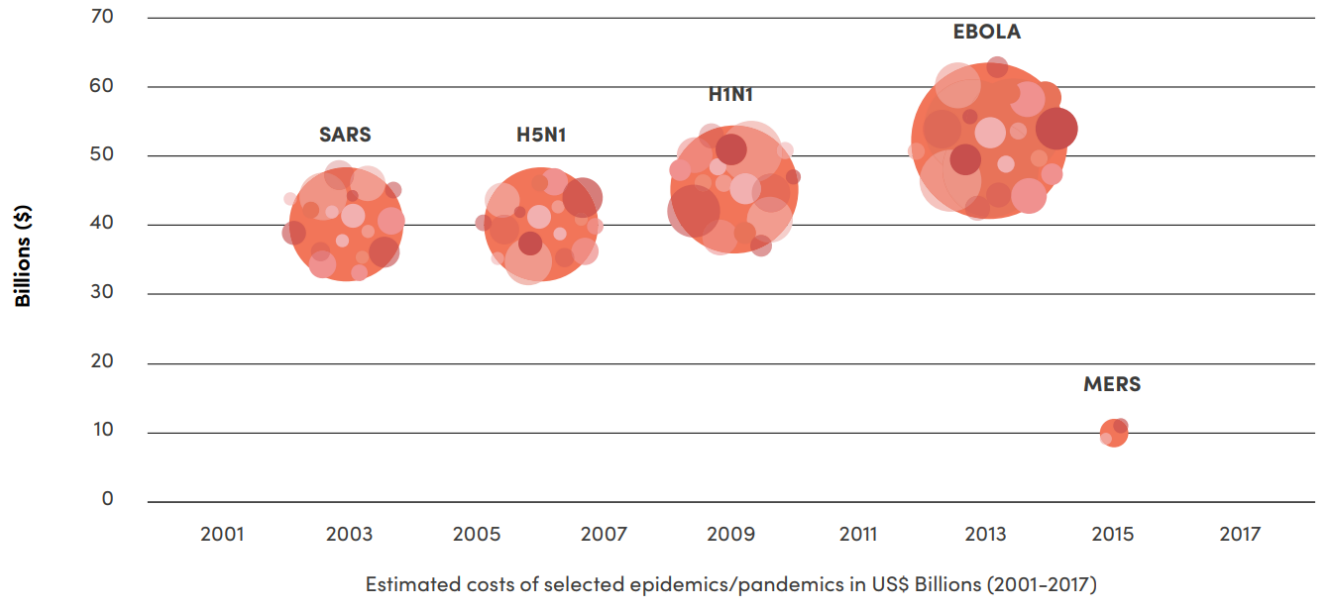
Determining the Tangible Conclusions that Can Be Drawn from COVID-19 Case Data for Strategic Supply Chain Decisions

Theodore Endresen （安多）

Significance

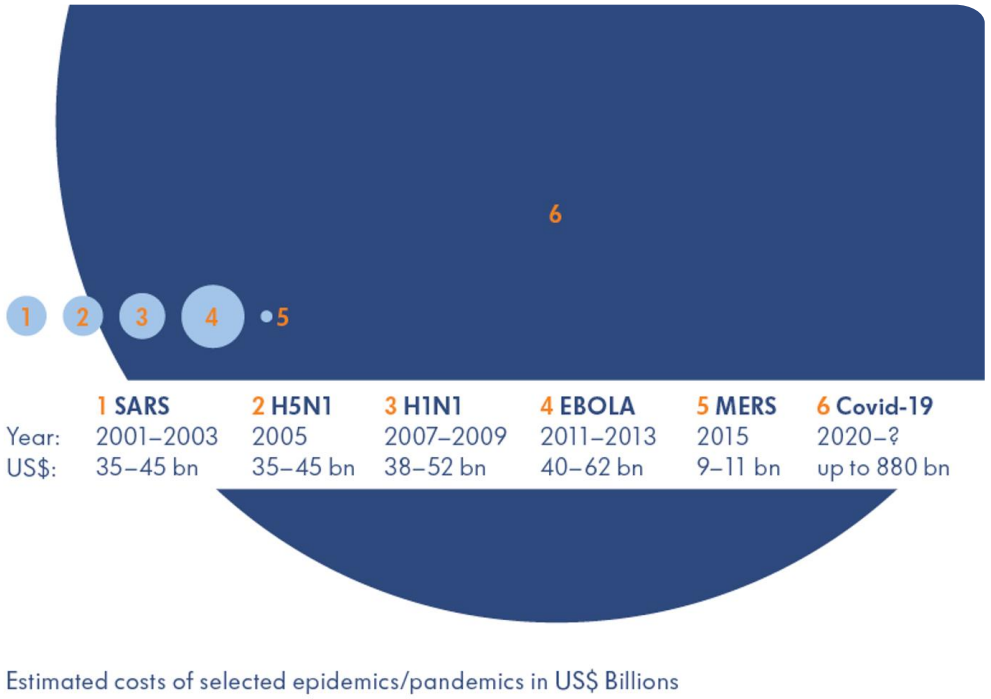
Why my research is important and would be valuable to pursue?



Significance

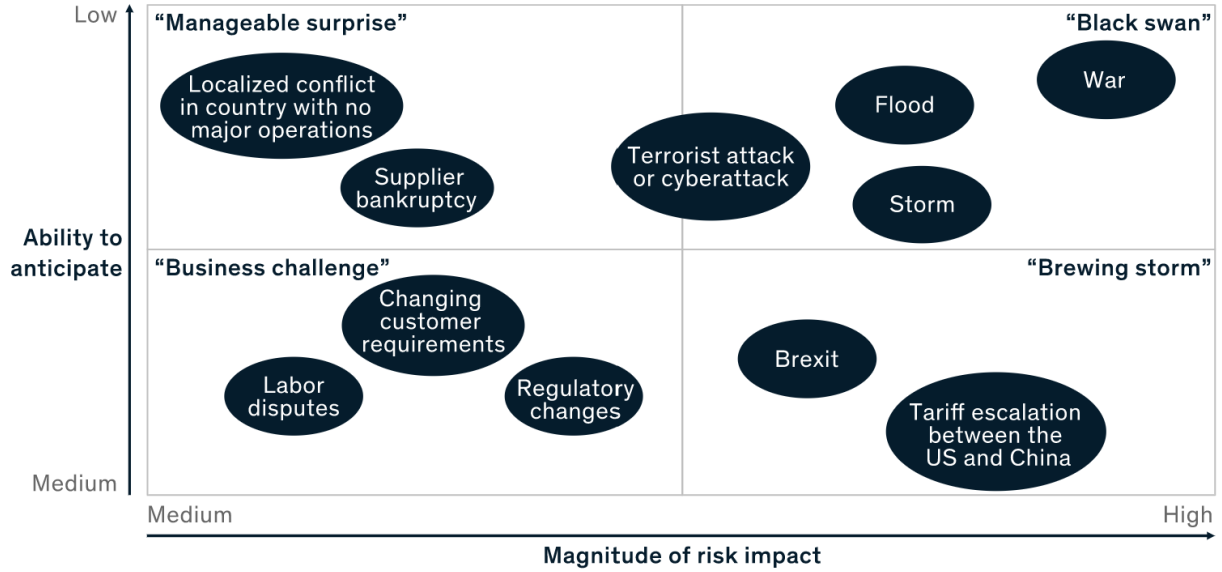
Pandemics prior to COVID-19 have extracted a massive economic impact on the global economy [1]

Significance



Estimated costs of selected epidemics/pandemics in US$ Billions

However, the economic impact of these prior pandemics pale in comparison to the estimated economic impact of COVID-19, which has estimated to be $880 billion in 2020 [2] and $12 trillion by 2040 [3].



Significance

In addition to current *Black Swan* supply chain risk strategies (i.e. wars, floods, storms, and terrorist attacks) pandemic preparedness will surely, and needs to be, considered in the future [4].

Significance

There exists a wealth of data documenting COVID-19 case, death, and testing numbers.

The question(s) companies will face is how to fairly and accurately parse through this data in order to draw information about a country's or region's pandemic preparedness from this data.

Research Objective

What I hope to accomplish through my research?

Research Objective

Overall Objective:

To sift through, and make sense of, the current COVID-19 case, death, and testing data in order to determine what conclusions can be drawn from it.

To do so, a list of metrics that captures the results of a country's COVID-19 response that can be drawn from the COVID-19 data must be created and calculated.

Research Objective

Then, these calculated metrics can be compared to economic, healthcare, and press independence data in order to try and discover and underlying trends.

Additionally, these calculated metrics can be compared to John's Hopkins Global Health Security (GHS) Rankings released in October 2019 [5], which are a measure of estimated pandemic preparedness.

Research Objective

Finally, based on the analyses, I should be able to draw tangible conclusions that determine the effectiveness of a country's COVID-19 response, what other prior known indicators played a significant role in predicting a countries observed COVID-19 response, and how accurate John's Hopkins 2019 GHS Rankings were.

Methodology

How I propose I should conduct my research?

Methodology – Date Sources

We need to following data for every country in order to conduct this research:

- Total Deaths
- Total Cases
- Total Tests

From this data, we should be able to extrapolate all necessary metrics to conduct our analyses

Methodology – COVID-19 Date Sources

The COVID-19 Case and Death data can be obtained through Johns Hopkins' Whiting School of Engineering's Center for Systems Science and Engineering public Github page [6].

The COVID-19 Testing data can be obtained through Our World in Data's public Github page [7].

Both these data sources have been vigorously tracking and publishing their data sources, so they can be appropriately used for our research.

Methodology – Economic Comparison

In order to do the economic comparison, we can utilize economic data from the International Monetary Fund's (IMF) April 2020 World Economic Outlook Report [8].

From this data, we will most likely look at the following data

- GDP Per Capita
- Unemployment Rate
- Government Finance

Methodology – Healthcare Comparison

In order to do the healthcare comparison, we can utilize healthcare ranking data from the World Health Organization (WHO) report titled *Measuring Overall Health System Performance For 191 Countries* [9].

This report created a composite score that weighted every country's healthcare's efficiency versus their quality and equity.

Methodology – Press Independece Comparison

In order to the press independence comparison, we can utilize the data from the 2020 World Press Freedom Index Rankings published by Reporters Without Borders [10]

In addition to ranking each country, the report also assigned each country a score which they have also correlated to a categorical rating.

Methodology – John's Hopkins Pandemic Preparedness Comparison

The Johns Hopkins GHS Rankings [5] pose an interesting area of study because they were released in October 2019. Thus, they give us a great view of what pandemic preparedness was viewed to be before the COVID-19 outbreak.

The reports ranks every country and provides a score across six dimensions of pandemic preparedness (prevent, detect, respond, health, norms, and risk) as well as an aggregated overall ranking.

Methodology – University of Oxford Stringency Index

Researchers at the University of Oxford have put together a Stringency Index that "systematically collects information on several different common policy responses that governments have taken to respond to the pandemic on 17 indicators such as school closures and travel restrictions" [11].

This data can also be used to leverage the COVID-19 metrics I calculate and the John's Hopkins Pandemic Preparedness Metrics.

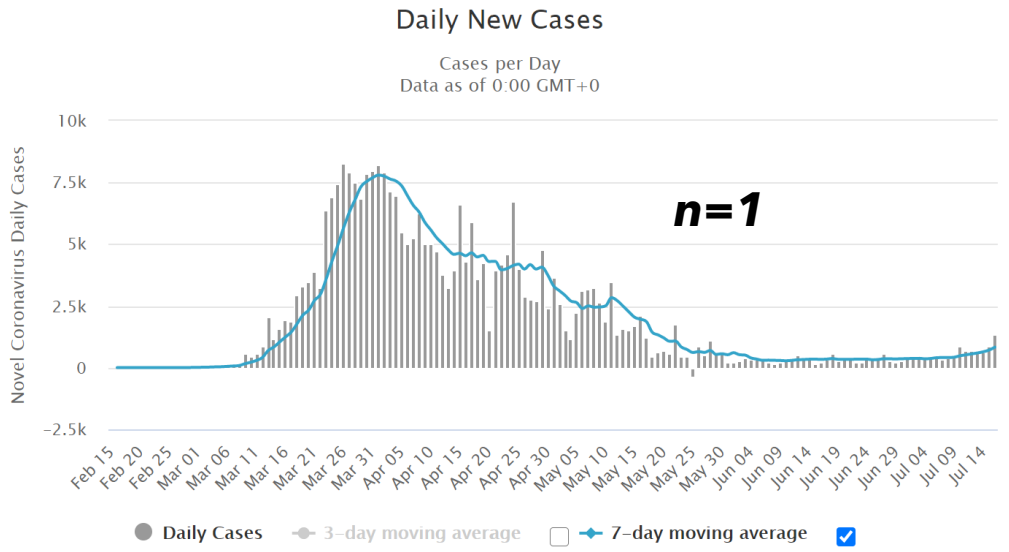
Methodology – Metrics

Based on the COVID-19 Cases, Deaths, and Testing data, we must put together a list of metrics that can be used to measure the results of a country's COVID-19 pandemic response.

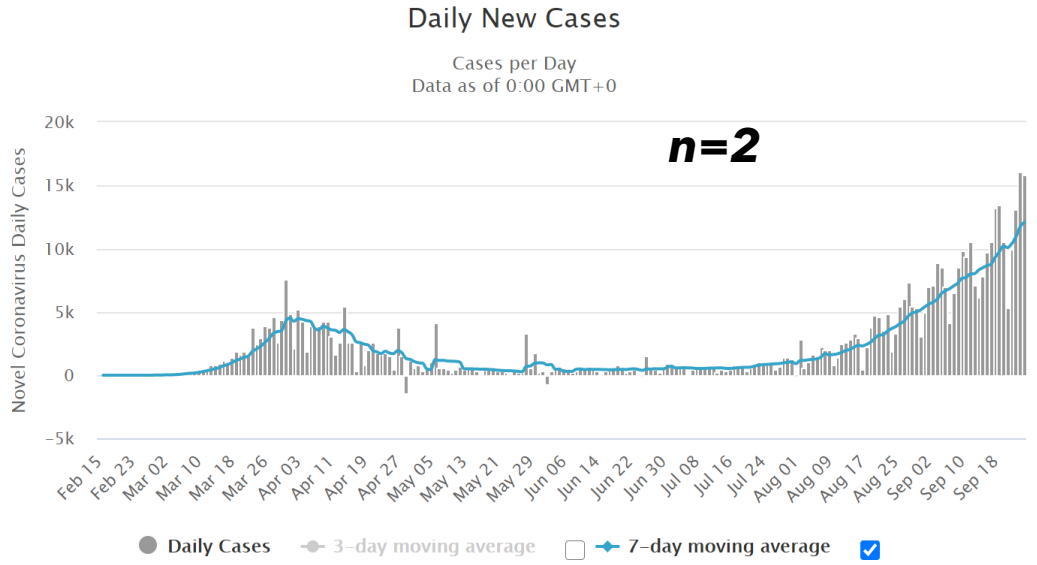
Methodology – Metrics: Number of Outbreaks

This metric would seek to explain the amount of outbreaks, *n*, that a single country experienced.

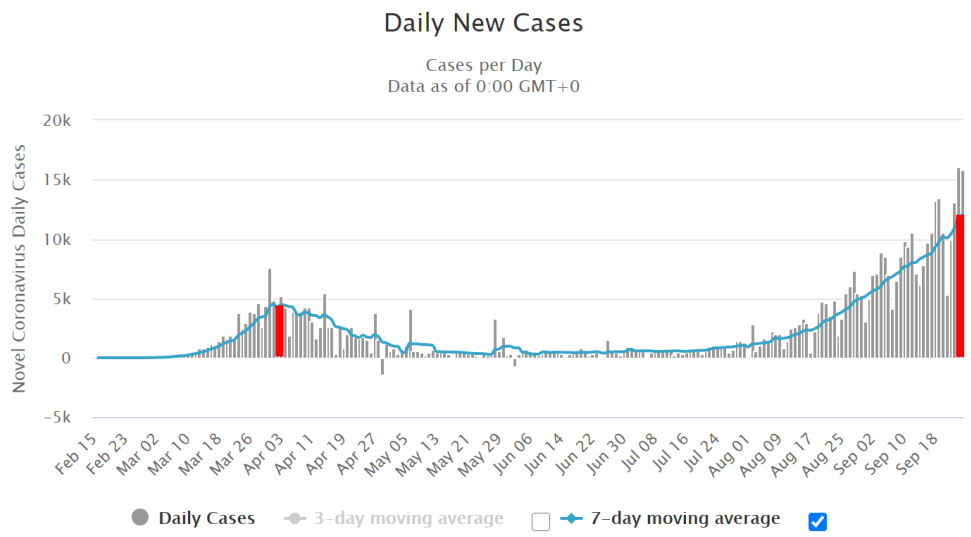
Daily New Cases in Spain

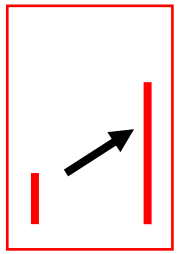


Daily New Cases in France

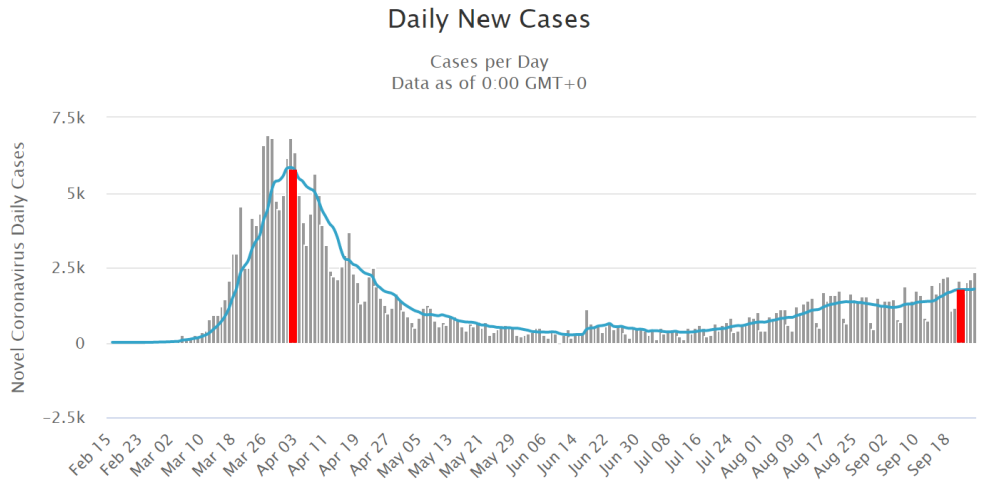


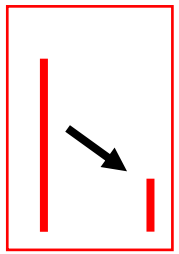
Daily New Cases in France





Daily New Cases in Germany

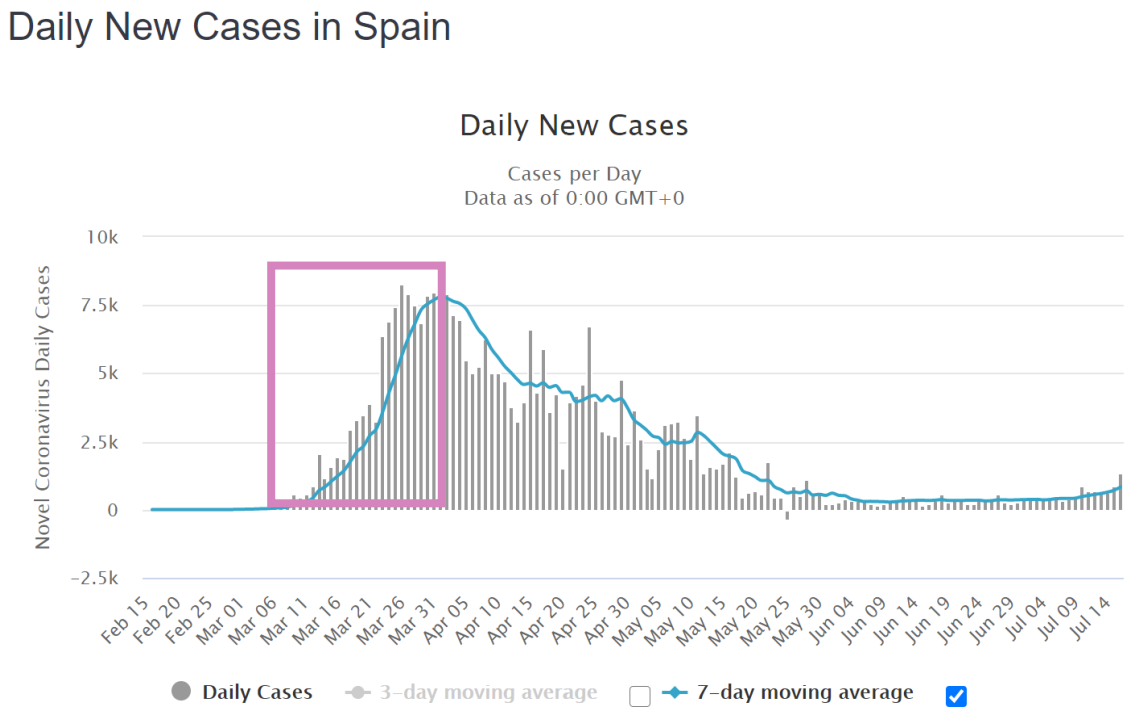




Methodology – Metrics: Severity of Subsequent Outbreaks

This metric would seek to explain the sustained response of a country's COVID-19 response. The lower the peak of subsequent outbreaks, the better the COVID-19 response.

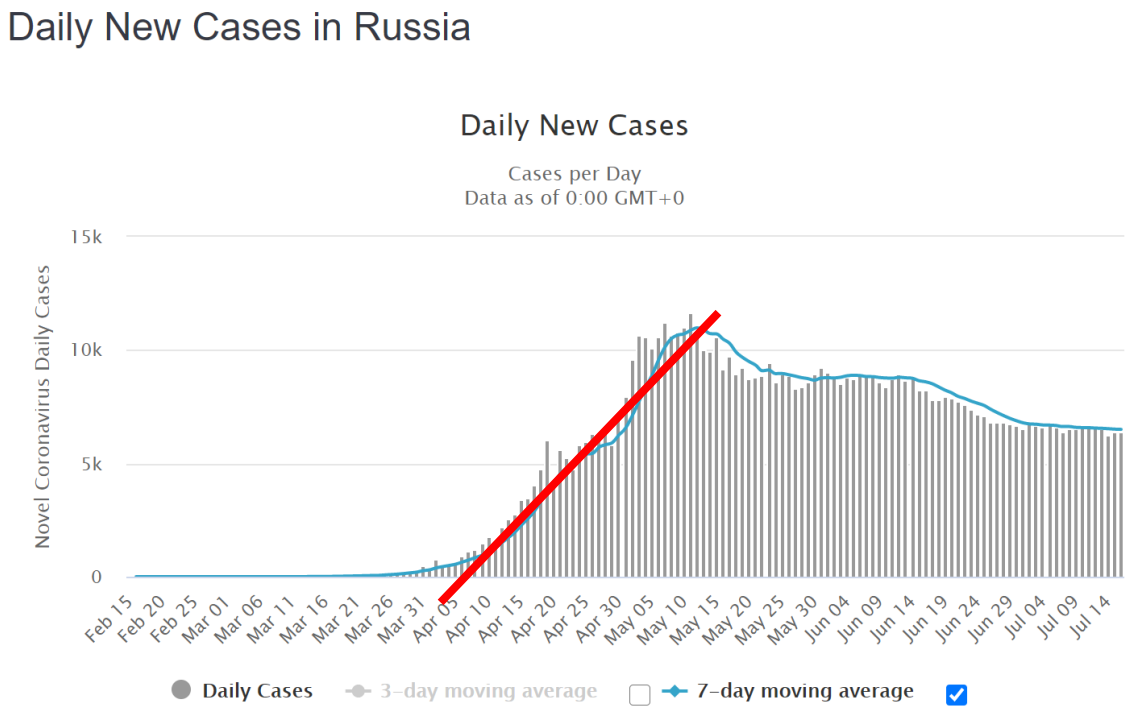
Methodology – Metrics: Length of Outbreak

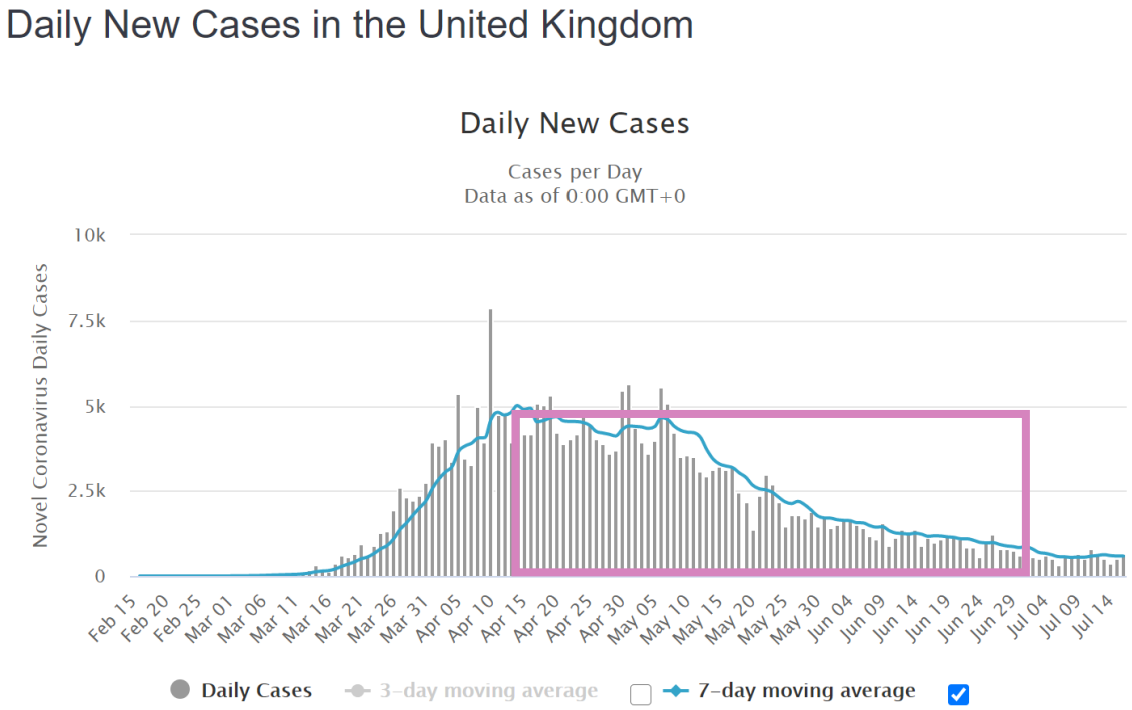


This metric would seek to explain the amount of time the pandemic within a country underwent uncontrolled growth during it's *nth* outbreak.

Methodology – Metrics: Rate of Outbreak

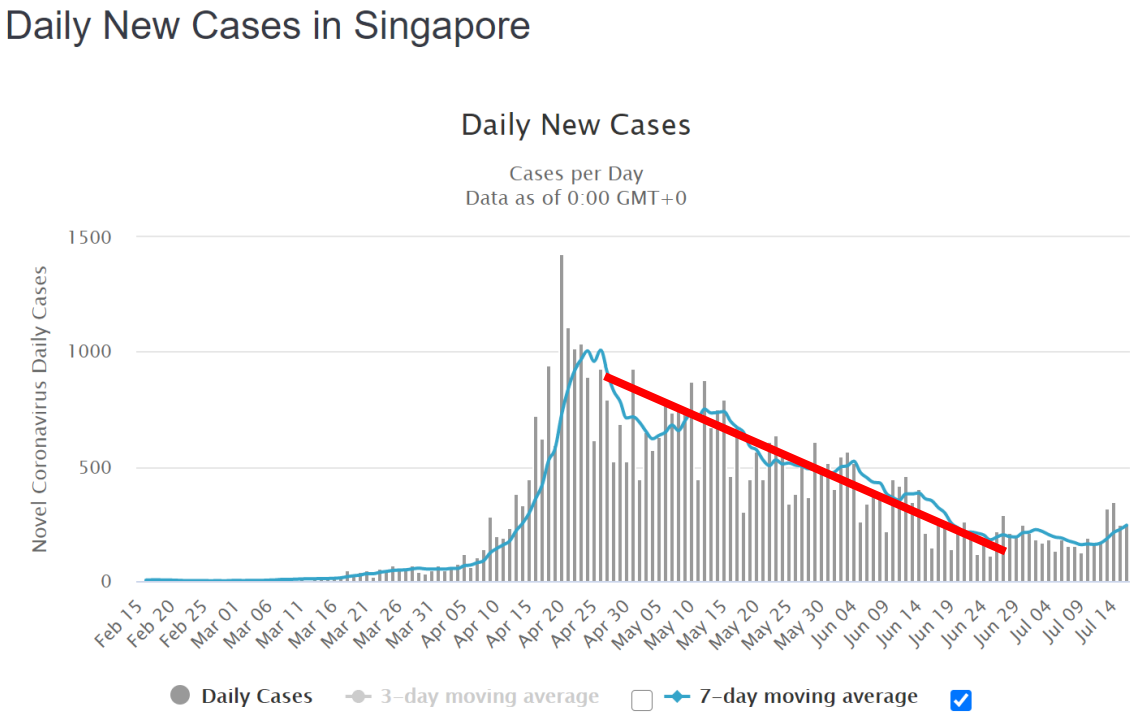
This metric would seek to explain the rate of uncontrolled growth a country faced during their nth outbreak.





Methodology – Metrics: Time to Submission

This metric would seek to explain the amount of time it took, after the pandemic within a country had stop experiencing uncontrolled growth, for a country to reduce its number of cases to where the pandemic could be considered submitted for it's *nth* outbreak.



Methodology – Metrics: Rate of Submission

This metric would seek to explain the rate at which a country got their pandemic under control, after the pandemic within a country had stop experiencing uncontrolled growth, to the point at which it could be considered under submission or began its subsequent *nth* outbreak.

Case to Death Ratio Average

This single data point would represent a country's death rate at the time of running the analysis. It is worthy to include because it may indirectly correlate to the quality of healthcare faculties in each country,

$$\frac{Total\ Deaths}{Total\ Cases}$$

Case to Death Ratio over Time

This metric would be a time series of the case to death ratio over time. This indicator is worthy to include because it may be an indicator of how effective a country's pandemic control measures were at protecting vulnerable populations.

Time	t=1	t=2	t=3	t=4
Ratio	$\frac{Deaths}{Cases}$	$\frac{Deaths}{Cases}$	$\frac{Deaths}{Cases}$	$\frac{Deaths}{Cases}$

Case to Tests Ratio Average

This single data point would represent the relationship between the number of tests done to the number of cases. This is worthy to include as it may indirectly relate to the quality of a country's healthcare system.

$$\frac{Total\ Cases}{Total\ Tests}$$

Case to Tests Ratio over Time

This metric would be a time series of the case to test ratio over time. This indicator is worthy to include because it may be an indicator of the quality of a country's healthcare system.

Time	t=1	t=2	t=3	t=4
Ratio	$\frac{Cases}{Test}$	$\frac{Cases}{Test}$	$\frac{Cases}{Test}$	$\frac{Cases}{Test}$

Testing Rate over Time

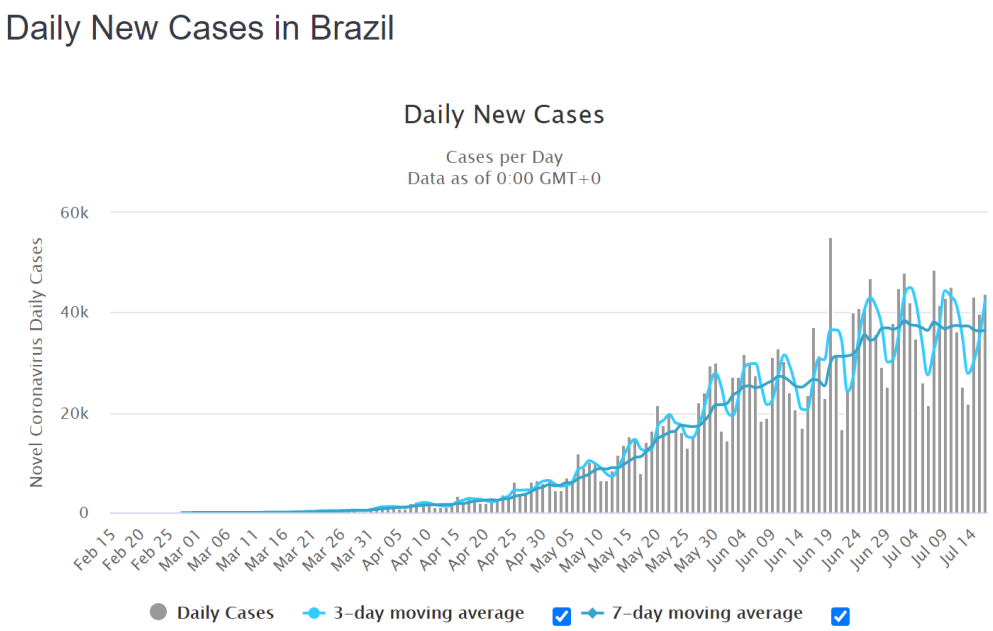
Time	t=1	t=2	t=3	t=4
# New Tests	$F(T = 1)$	$F(T = 1)$	$F(T = 1)$	$F(T = 1)$

This time series metric would show how many tests a country was able to do over time. This would be important to know as it would show how well a country was able to respond to the pandemic.

Known Challenges and Proposed Solutions

The challenges I know I must solve in order to conduct my research and how I think they can be sovled.

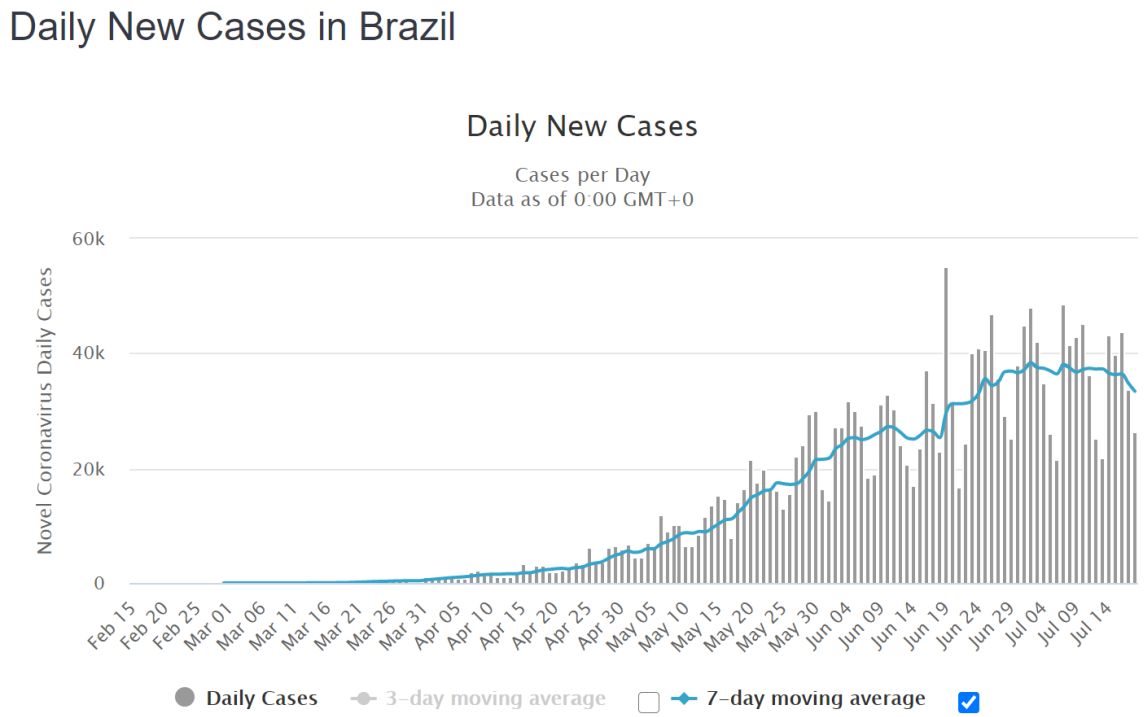
Periodicity of Data



Many country's data exhibits a great deal of weekly periodicity. Therefore, in order to make the data processing more accurate, a 7-10 day moving average should be applied.

How to Measure Cases, Deaths, and Tests

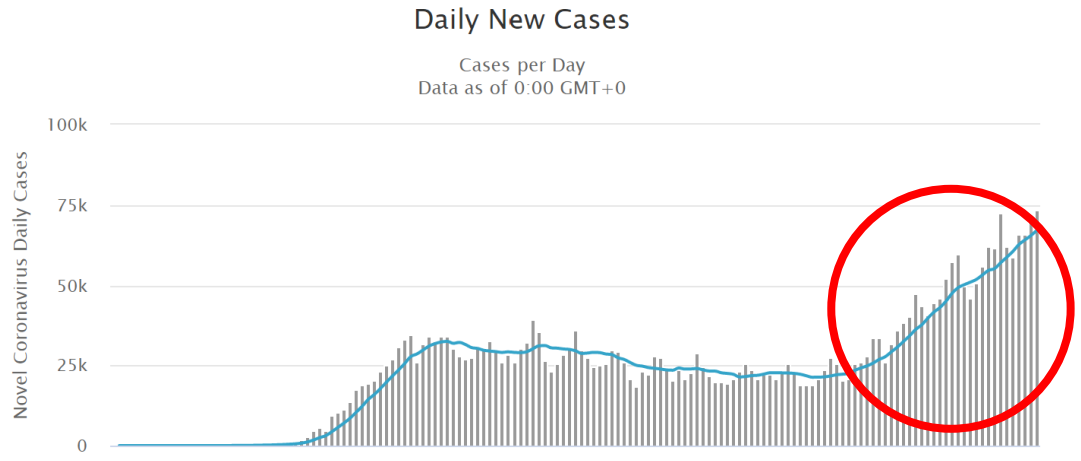
Countries have vastly different populations. Therefore, to make sure all countries are compared accurately, all data needs to be converted to indicants per 100,000 or 1,000,000 people.



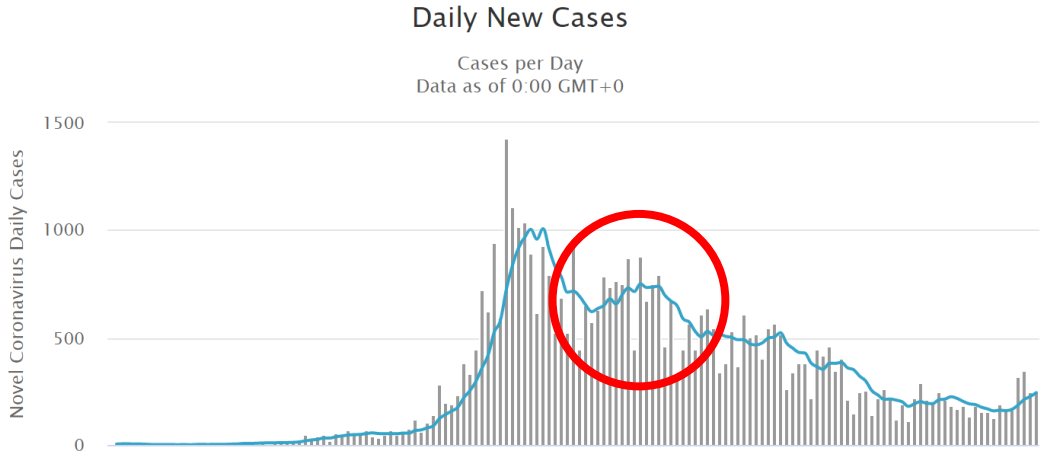
Describing and Smoothing Distributions

In order to obtain useful information from the COVID-19 case data, we must first smooth the data in order to remove the elements of weekly periodicity from the data. This can be achieved, as previously mentioned, through a moving average.

Daily New Cases in the United States



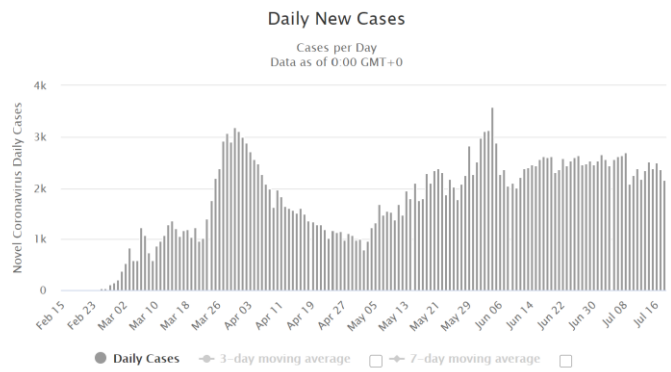
Daily New Cases in Singapore



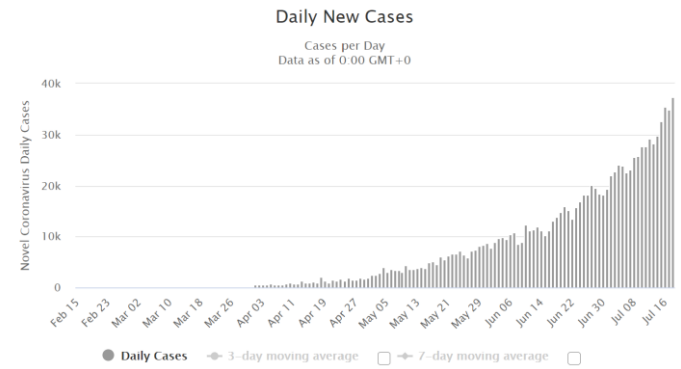
Describing and Smoothing Data

However, even by utilizing such smoothing methods, the data is not quite usable to determine the start and end points of different portions of curves. For example, although Signapore saw an increase in cases, I would not consider it a second outbreak like we see in the US Data.

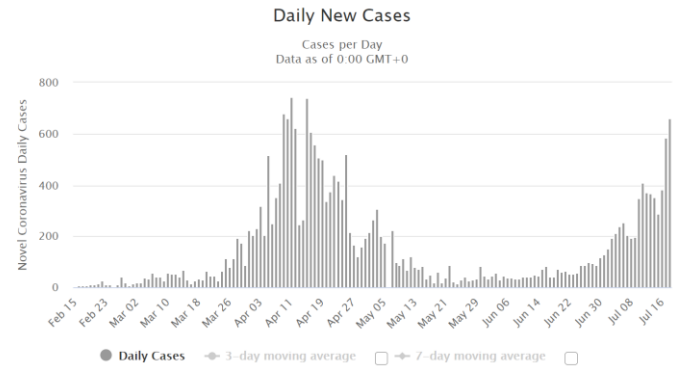
Daily New Cases in Iran



Daily New Cases in India



Daily New Cases in Japan



Describing and Smoothing Data

Although we could go through every country's data and manually choose start and end points, this would be a worst-case scenario. However, due to the sheer number of countries we need to analyze, creating a list of rules that would encapsulate every country's unique data would be cumbersome.

Describing and Smoothing Data

To solve this issue, I think we can utilize the following python libraries to create curve fitted functions from which we can determine the first derivatives to determine the start and end points of the different outbreaks.

- curve_fit
  - Utilizes optimization to determine the coefficients of a pre-input non-linear function
- f_solve
  - Utilizes optimization to determine the roots of an inputted nonlinear function

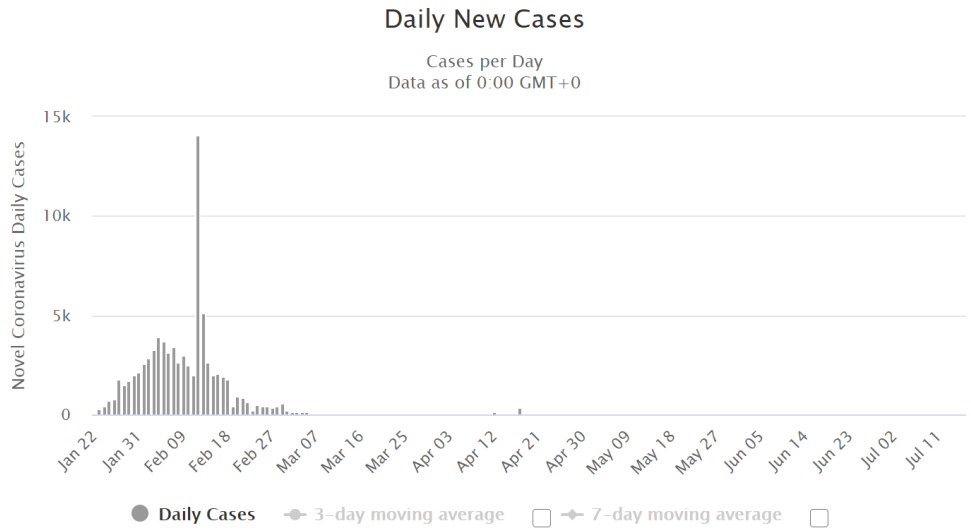
Describing and Smoothing Data

By feeding different functions (i.e. quadratic, cubic, quartic, exponential, Poisson, etc.) to curve_fit to be fitted to the data and selecting the result with the best R^2 value, we can systemically and repeatably get a smoothened function for which to determine an outbreak's start and end points.

Dealing with Extraneous Data Points

Some countries have retroactively increased or decreased daily incidents to account for counting or other discrepancies. This data either needs to be completely removed, to equally distributed across the prior two weeks.

Daily New Cases in China



Analysis Methods

What methods will be used to determine if relationships exist between my metrics and existing rankings/metrics.

Trend Analysis

For results where we seek to identify if a trend occurs, we can fit a trendline (linear or nonlinear) and use the resulting R^2 value to determine the significance of the resulting correlation.

ANOVA Analysis

For the analysis of comparison results, we can use ANOVA Analysis to determine if a significant difference between the data occurs via T-Tests or other statistical tests.

Additional Work – My Pandemic Response Rankings

If time allows, as an addendum to my research, it may be interesting to create my own rankings of a country's pandemic response based off the COVID-19 metrics I calculate by utilizing ratio scaling methods.

Sources

[1] https://apps.who.int/gpmb/assets/annual_report/GPMB_annualreport_2019.pdf
[2] https://www.medelahealthcare.com/en-GB/insights/fighting-the-pandemic-takes-preparation
[3] https://www.mckinsey.com/business-functions/risk/our-insights/covid-19-implications-for-business#
[4] https://www.mckinsey.com/business-functions/operations/our-insights/supply-chain-risk-management-is-back
[5] https://www.ghsindex.org/about/
[6] https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data
[7] https://github.com/owid/covid-19-data/tree/master/public/data
[8] https://www.imf.org/external/pubs/ft/weo/2020/01/weodata/index.aspx
[9] https://www.who.int/healthinfo/paper30.pdf?ua=1
[10] https://rsf.org/en/ranking
[11] https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker