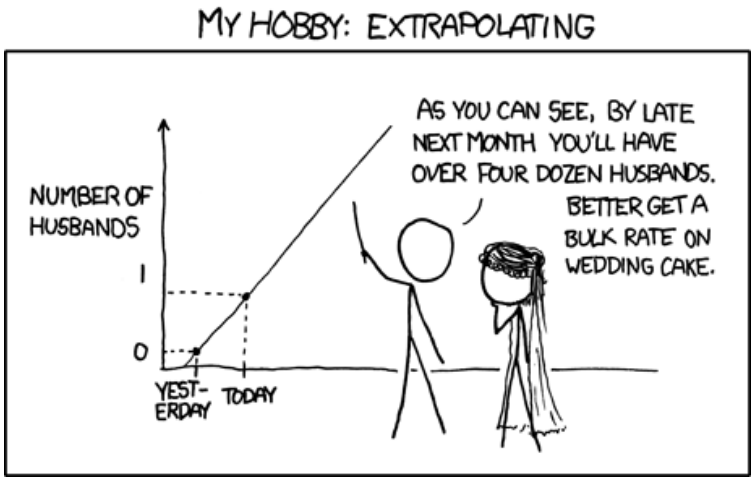
Epidemic Modeling 101
Why your CoVID-19 exponential fits are wrong

medium.com/data-for-science/epidemic-modeling-101-or-why-your-covid19-exponential-fits-are-wrong-97aa50c55f8

Bruno Gonçalves, Medium, April 5, 2020

1. Why your CoVID-19 exponential fits are wrong

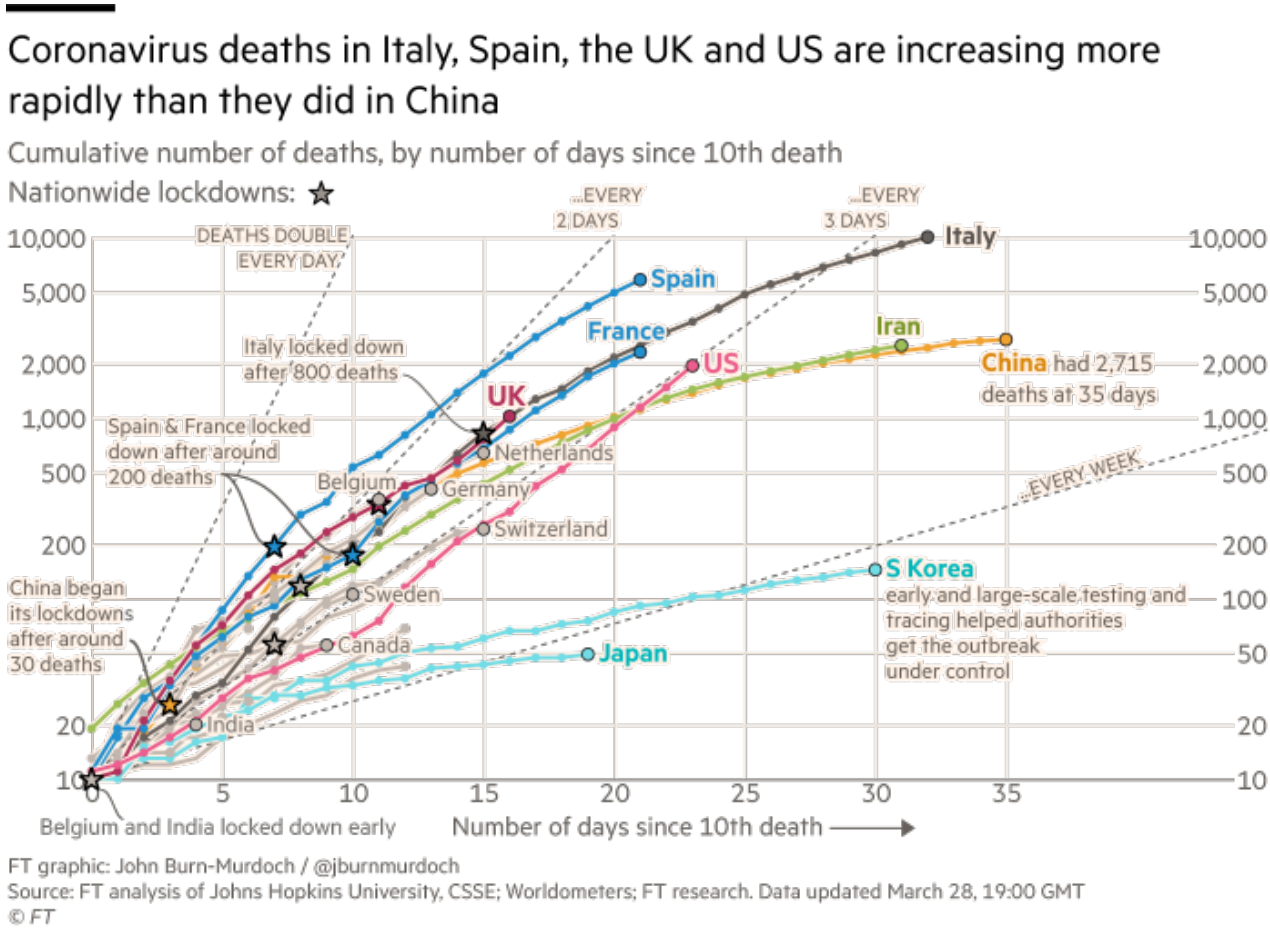
Over the past few weeks, a terrible affliction has been spreading across the world. Otherwise healthy and productive members of society have been infected with this devastating illness that causes them to fire up Excel, Python or R and start extrapolating the latest numbers of confirmed CoVID-19 cases in their town, state, country or even the entire world!



All joking aside, the severity of the current SARS-CoV-2 epidemic is undeniable and it is only natural that people will deal with the added stress in their lives (and extra free time due to lockdown procedures) in various ways.

A particularly afflicted demographics has been my own, that of Physicists, resulting in the rise of a small cottage industry of blog posts, LinkedIn publications and even arXiv papers with their best attempts at modeling the spread of the disease, with little or no understanding of dynamics underlying epidemic spreading. Once again, the immortal words of Simon DeDeo have been proven true:

Invariably, our fearless followers of John Snow (not the one you're thinking of) end up with some variation of this plot comparing the cumulative number of cases or deaths in various countries as a function of time with a straight exponential growth rate.



Financial Times, March 29, 2020

Extrapolation to unrealistic numbers, forecasts about when a country might overtake another, considerations about the success or failure of containment measures and various other shenanigans ensue.

Bringing order to a chaotic world has always been the driving force of Human progress and it can be argued that this is simply its latest incarnation: The Numerati trying to use their modeling and Data Science skills to make sense of the world around them. A trend that has led in recent years to impressive progress in Machine Learning, Artificial Intelligence, and Data Science. Unfortunately, while there are good reasons to expect the early stages of epidemic spread to be exponential, there are many practical factors conspiring against the efficacy of simple curve fitting and a little background knowledge about traditional epidemic modeling can go a long way.

What follows is my personal perspective, as an individual with some real world experience in epidemic modeling during previous pandemics and shouldn't reflect on any group or institution I might be affiliated with.

Compartmental Models

Mathematical modeling in Epidemiology has a long and rich history, dating as far back as the 1920s with Kermack–McKendrick theory. The basic idea is deceptively simple: we can divide the population into different compartments representing the different stages of the disease and use the relative size of each compartment to model how the numbers evolve in time.

In the discussion below, I introduce several simple models and scenarios to help illustrate the issues with simply trying to do curve fitting on the empirical numbers. You can find the notebook I wrote to implement the models and generate the figures over at the GitHub repository I made specifically for this post:

DataForScience/Epidemiology101

Contribute to DataForScience/Epidemiology101 development by creating an account on GitHub.

github.com

SI Model

Let's start by taking a look at the simplest possible epidemic model: The *Susceptible-Infected* model. Here we split our population into two compartments, the healthy compartment (usually referred to as ***Susceptible***) and the ***Infectious*** compartment. The dynamics is also simple, when a healthy person comes in contact with an infectious person s/he becomes infected with a given probability. And, in this simple example, when you are infected, you remain infected forever. Mathematically, this is often written as:

which is just a fancy way of saying that the loss in the number of healthy people is the same as the gain in the ranks of the infected. More specifically:

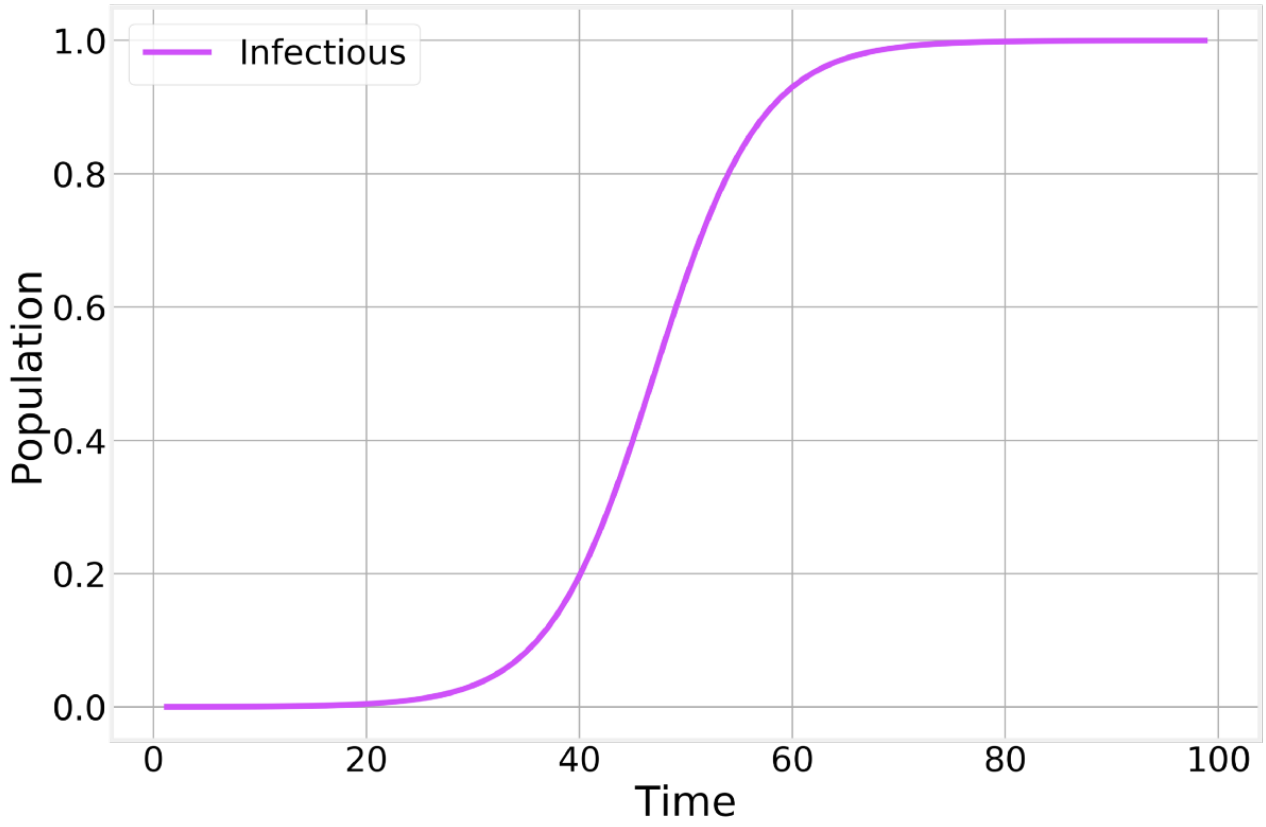
- N is simply the total population size
- β is the rate of infection
- It/N is the fraction of infected people and it represents the probability that a susceptible person will encounter an infected one.

$$\frac{\partial}{\partial t} S_t = -\beta S_t \frac{I_t}{N}$$

$$\frac{\partial}{\partial t} I_t = +\beta S_t \frac{I_t}{N}$$

Mathematical description fo the Susceptible-Infected model

Not surprisingly, this is not a very interesting model: given enough time everyone becomes infected:



Infectious fraction of the total population as a function of time.

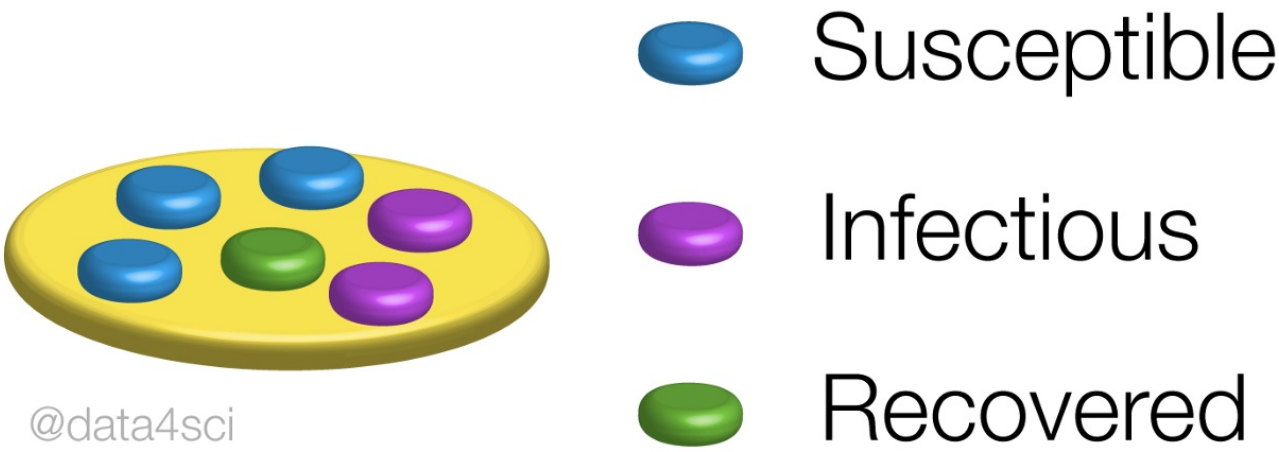
This simple model considers only one way to transition between compartments: From S to I through the interaction (contact) between S and I. A compact way to represent this is:

$$S + I \xrightarrow{\beta} I$$

The transition in the SI model

SIR Model

More realistic epidemic models can be developed by adding further compartments and transitions. The most common such model is the *Susceptible-Infectious-Recovered* model:



SIR Model

Here we have a new compartment, *Recovered*, that represents the people who have had the disease in the past and have since recovered, becoming immune. The presence of *Recovered* slowly reduces the number of infectious individuals as they are allowed to recover.

In terms of transitions this can be written as:

Where the second line represents a spontaneous (non-interacting) transition from *Infectious* to *Recovered* at a fixed rate μ.

$$S + I \xrightarrow{\beta} I$$

$$I \xrightarrow{\mu} R$$

The transitions in the SIR model

Or, mathematically, as:

which makes it clear that the growth in the number of *Recovered* depends only on the current number of *Infectious* individuals. It should also be noted that this model implies a constant population size:

A similar expression could be written for the SI model as well.

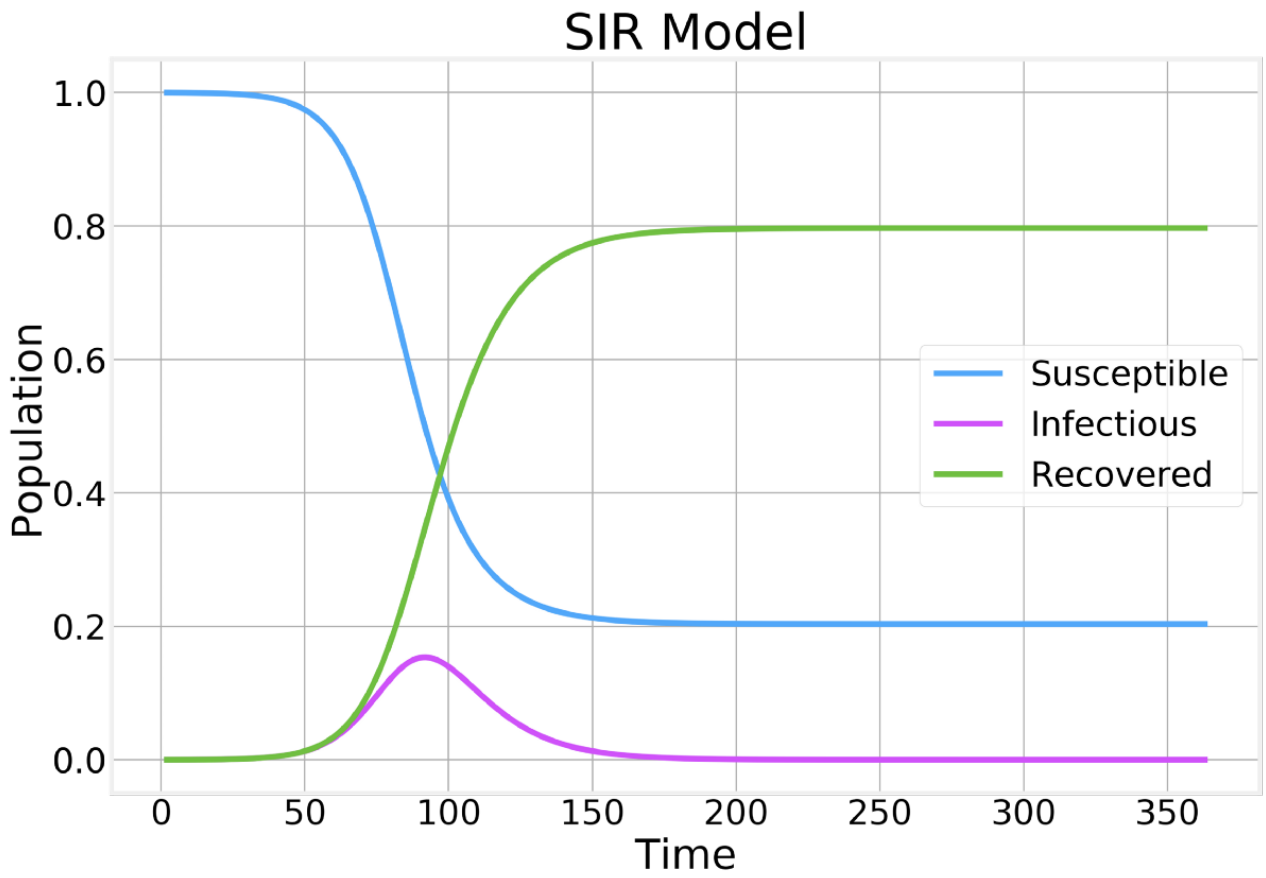
If we now integrate the full SIR model, we find:

$$\begin{aligned}
\frac{\partial}{\partial t} S_t &= -\beta S_t \frac{I_t}{N} \\
\frac{\partial}{\partial t} I_t &= +\beta S_t \frac{I_t}{N} - \mu I_t \\
\frac{\partial}{\partial t} R_t &= +\mu I_t
\end{aligned}$$

Susceptible-Infectious-Recovered model

$$\frac{\partial}{\partial t} S_t + \frac{\partial}{\partial t} I_t + \frac{\partial}{\partial t} R_t = 0$$

Fixed total population

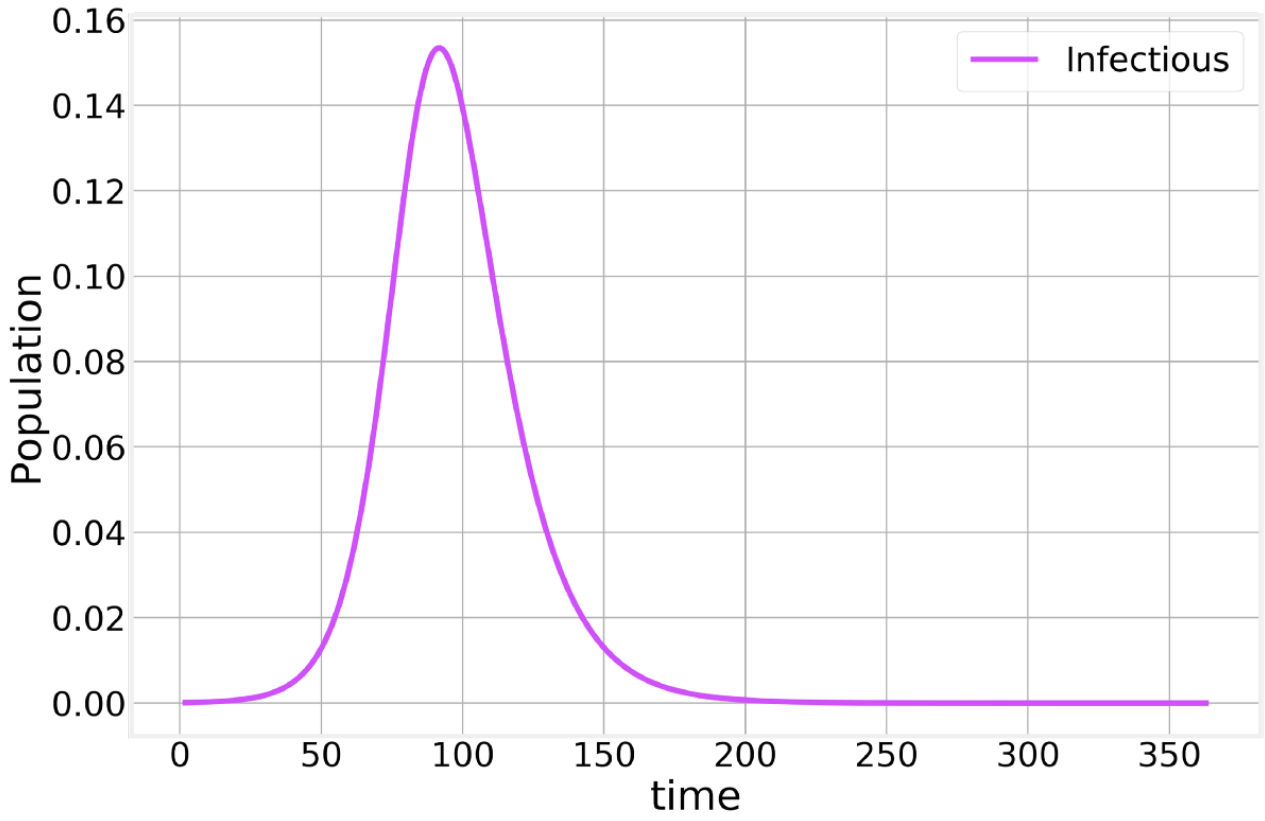


Fraction of the population in each compartment as a function of time

A few things should be noticed about this plot:

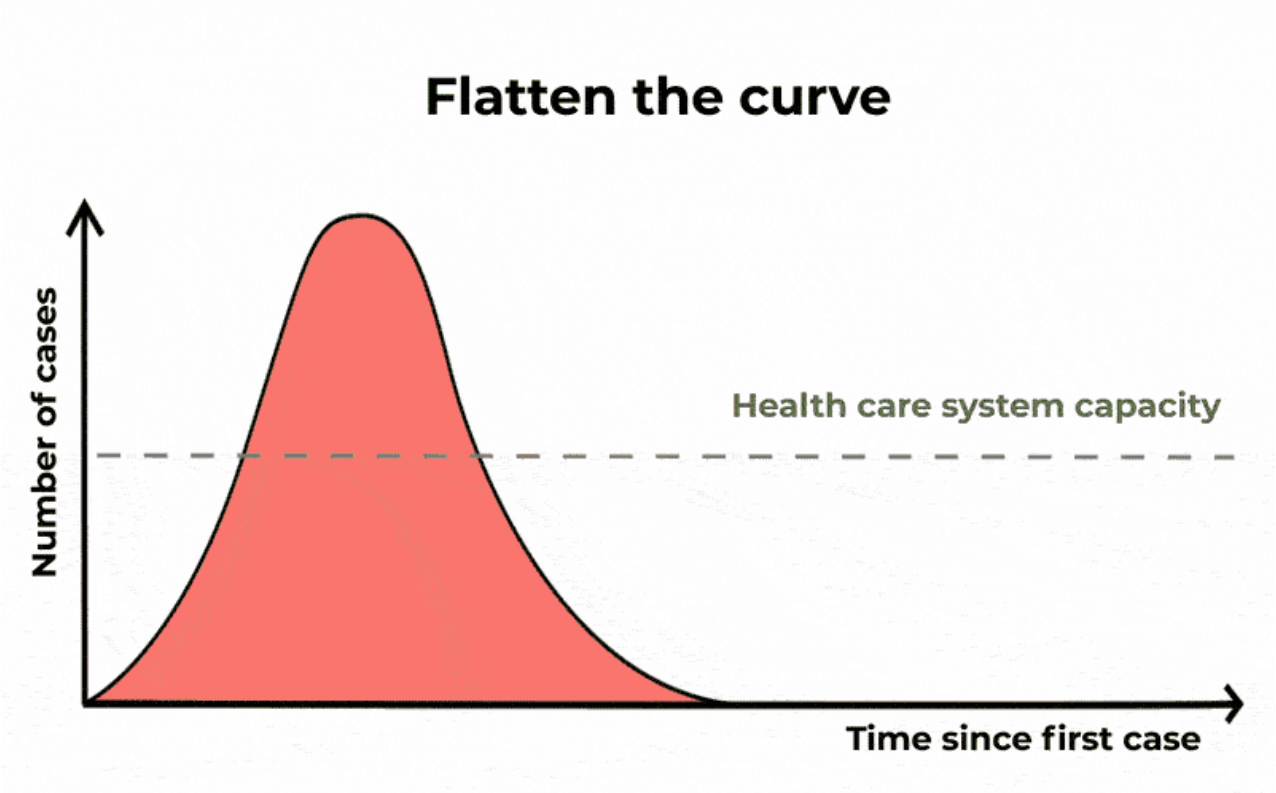
- The number of *Susceptible* individuals can only decrease
- The number of *Recovered* can only increase
- The number of *Infectious* individuals grows up to a certain point before reaching a peak and starting to decline.
- The majority of the population becomes infected and eventually recovers.

If we zoom in on just the behavior of the *Infectious* compartment, we find:



SIR Infectious compartment

Meaning that a significant fraction of the population can be infected at the same time, potentially causing (depending on the severity of the infection) the Healthcare system to become overwhelmed. When you hear about “flattening the curve” this is the curve that they are referring to.



The Conversation/CC BY ND

From the mathematical expression of the SIR model above, a few interesting results can be easily obtained. If we focus on the early days of the epidemic spreading, we can assume that the fraction of *Susceptible* individuals is still *~1* and find:

The exponential that everyone is trying to fit! Here,

$$I_t \approx I_0 e^{\mu(R_0-1)t}$$

is pronounced “R naught” and is the Basic Reproduction Number of the disease. This simple number defines whether or not we have an epidemic. If $R_o<1$ the disease dies off, otherwise, it grows *exponentially*!

One intuitive way of interpreting the R_o is as the average number of new infections produced by a single infectious individual. If a person is able to spread the disease to at least another one before recovering, then the epidemic can continue, otherwise, it dies off.

This is what we need to determine and it depends on many different factors **that are characteristic of the virus**, as Kate Winslet eloquently put it in the 2011 movie, Contagion.

The current best estimates of the R_o value for SARS-CoV-2, the coronavirus that causes CoVID-19 is around 2.5.

The value of R_o also plays a fundamental role in determining the course of the epidemic. If we consider the second equation describing the SIR model:

We find that the derivative of the number of infectious becomes negative whenever:

$$\frac{\partial}{\partial t} I_t = +\beta S_t \frac{I_t}{N} - \mu I_t$$

This is the point at which we have reached the peak and the epidemic starts dying off. This is the point at which the population starts having enough of what is known as *Herd immunity* for the disease to be unable to spread further. Whenever vaccines are available, vaccination programs are designed to help the population reach herd immunity without having to get a significant fraction of the population infected.

$$S_t < \frac{N}{R_0}$$

R_o also determines the final fraction of the entire population that will be unaffected by the disease:

Where *S_infinity* refers to the total fraction of healthy (and never infected) individual after the epidemic has had time to follow its course completely. This expression

$$S_\infty = e^{-R_0(1-S_\infty)}$$

isn't amenable to closed form solution, but can be used to numerically estimate the value of *S_infinity*. The SIR figure above was generated by using $R_o=2$ and we see that *S_infinity~0.2* which can be easily verified by plugging these numbers in this expression.

Practical considerations

So far, our analysis of epidemic models has focused on the ideal scenario which seems to justify the approach of fitting exponential curves as a simple way of trying to forecast the course of the epidemic. Unfortunately, the real world is significantly more complex in a variety of ways.

Asymptomatic and mildly infectious cases

One of the limitations of the approach described so far is that it makes a few unrealistic assumptions:

- There is **no incubation or latent period**. An incubation period delays the entire epidemic timeline. An issue that is not significant for our purposes here.
- There is a **single type of infectious** individual. In the real world, different immune systems respond differently to the virus resulting in some people being completely asymptomatic (no symptoms whatsoever) and mildly infectious cases. In the case of CoVID-19 the number of asymptomatic cases is thought to be 40% or higher.

Both of these difficulties can be addressed by adding new compartments and transitions to our basic SIR model without much difficulty. However, they pose significant challenges when dealing with the official published numbers.

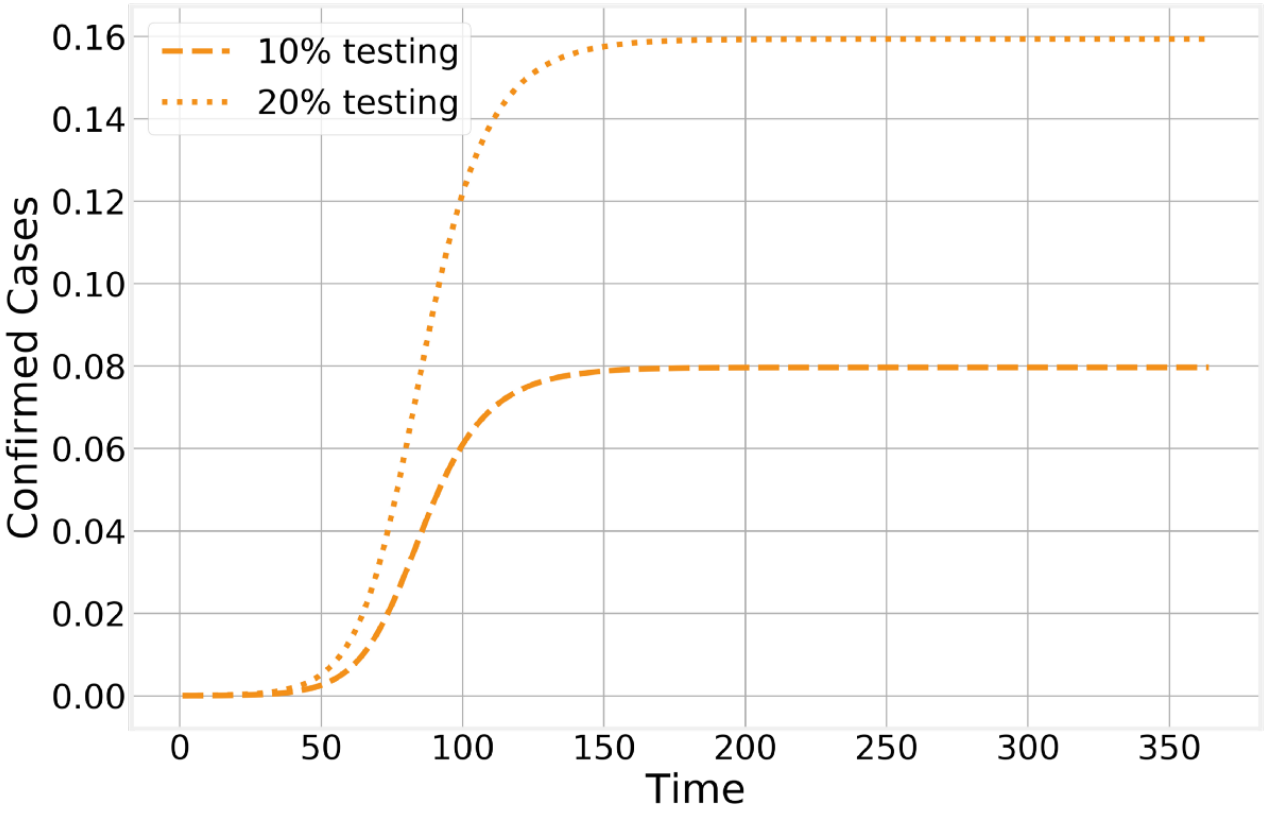
In the early days of the epidemic, only the more severe cases (non-asymptomatic and non-mild) cases get sick enough to search medical help and be officially diagnosed. This naturally leads to a delay in detection of the first cases in a given city or country and an over-estimation of the severity of the disease as more severe cases are more likely to die.

Published numbers are also typically cumulative, making the total numbers appear larger. A simple way of extracting a measure of the number of possible confirmed cases from our simple SIR model is to count how many people have been removed from the *Susceptible* compartment. Defining ϕ to be the fraction of infectious cases that do get tested, we have:

As a result, the numbers that get published depend directly on the fraction of cases that are severe enough to both lead to medical attention and be tested:

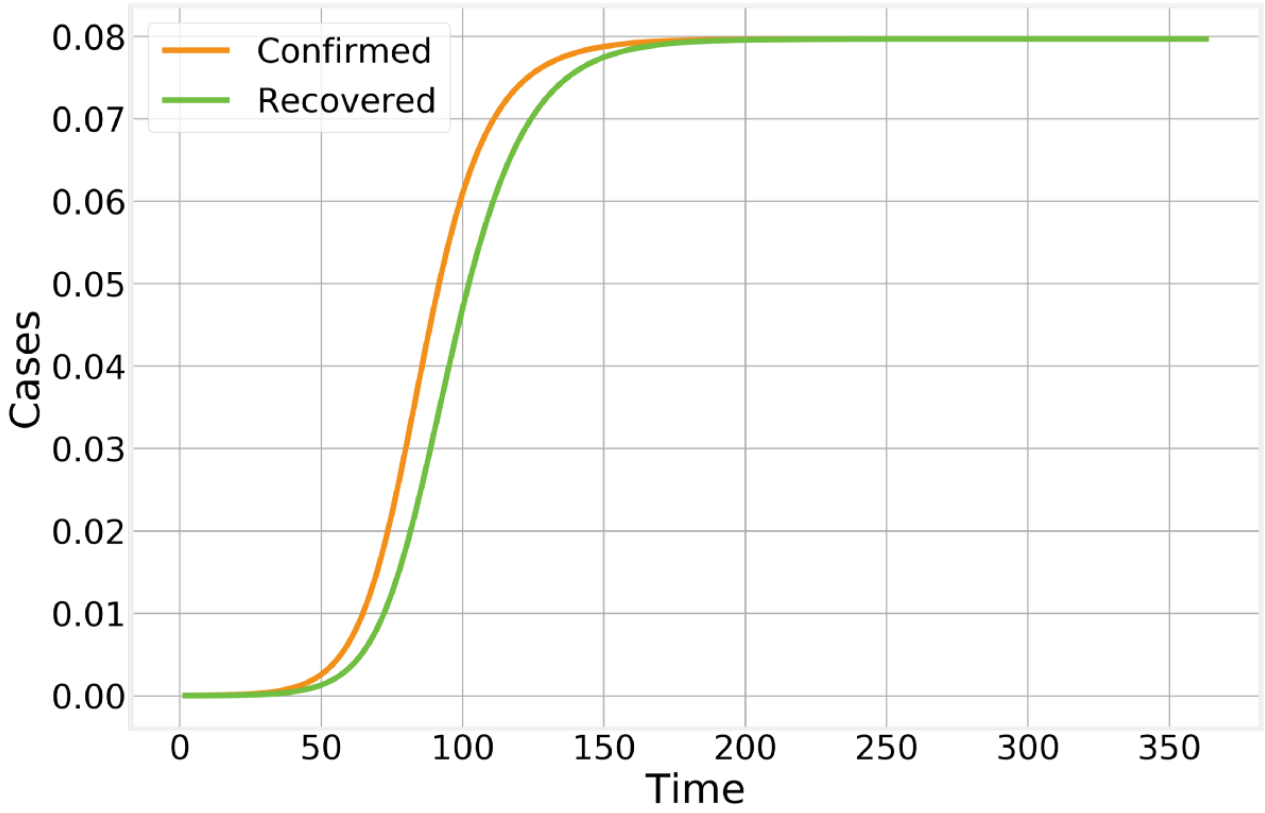
$$C_t = \phi\left(N - S_t\right)$$

Confirmed cases



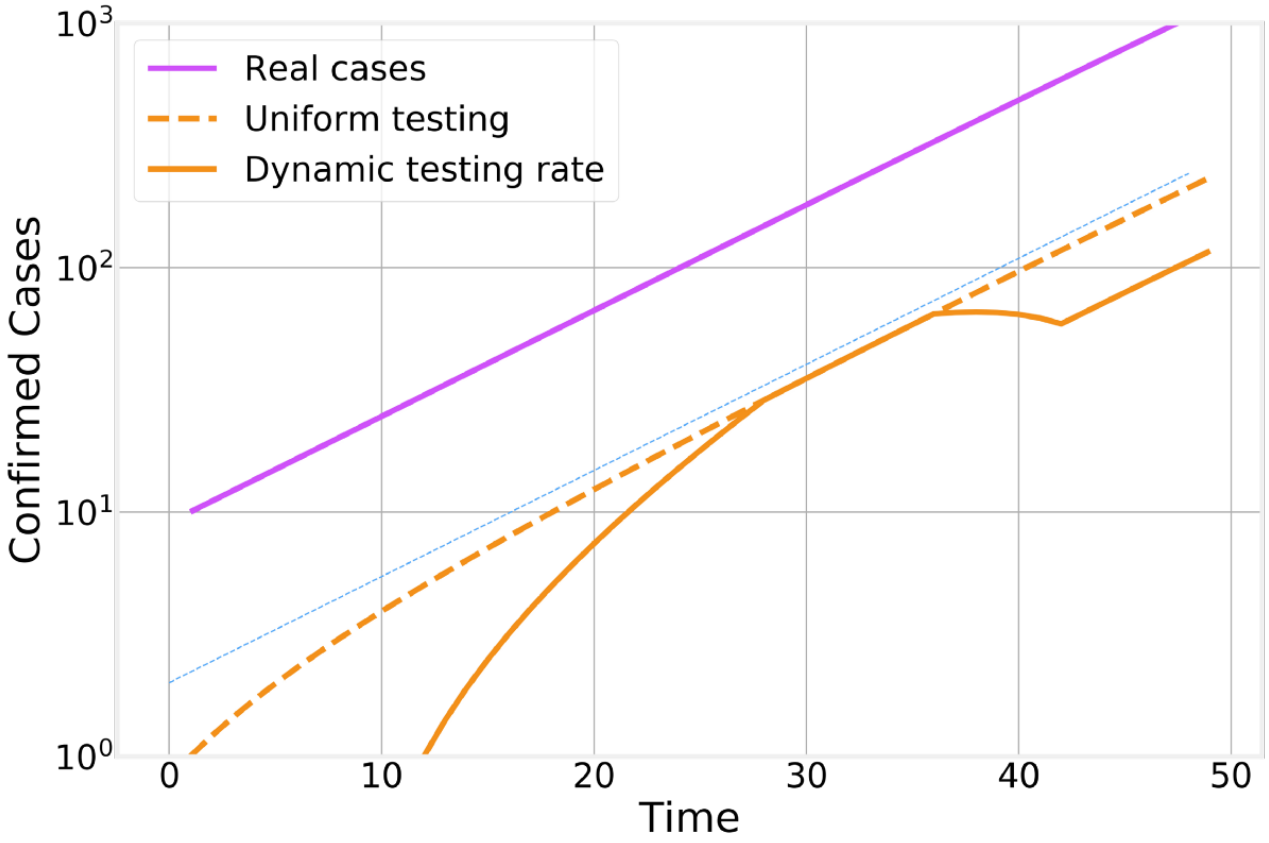
Confirmed cases in the SIR model

The number of (observed) *Recovered* individuals will then follow a similar trajectory, although with a f*ew days lag due to the natural time line of the disease:*



Observed recovered number of cases

Naturally, with novel diseases it takes time to develop and distribute accurate tests. If we further consider that the testing fraction ϕ is time dependent as well, then it is easy to see how a lot of the features observed in the time line of confirmed cases are caused by local policies and test availability:



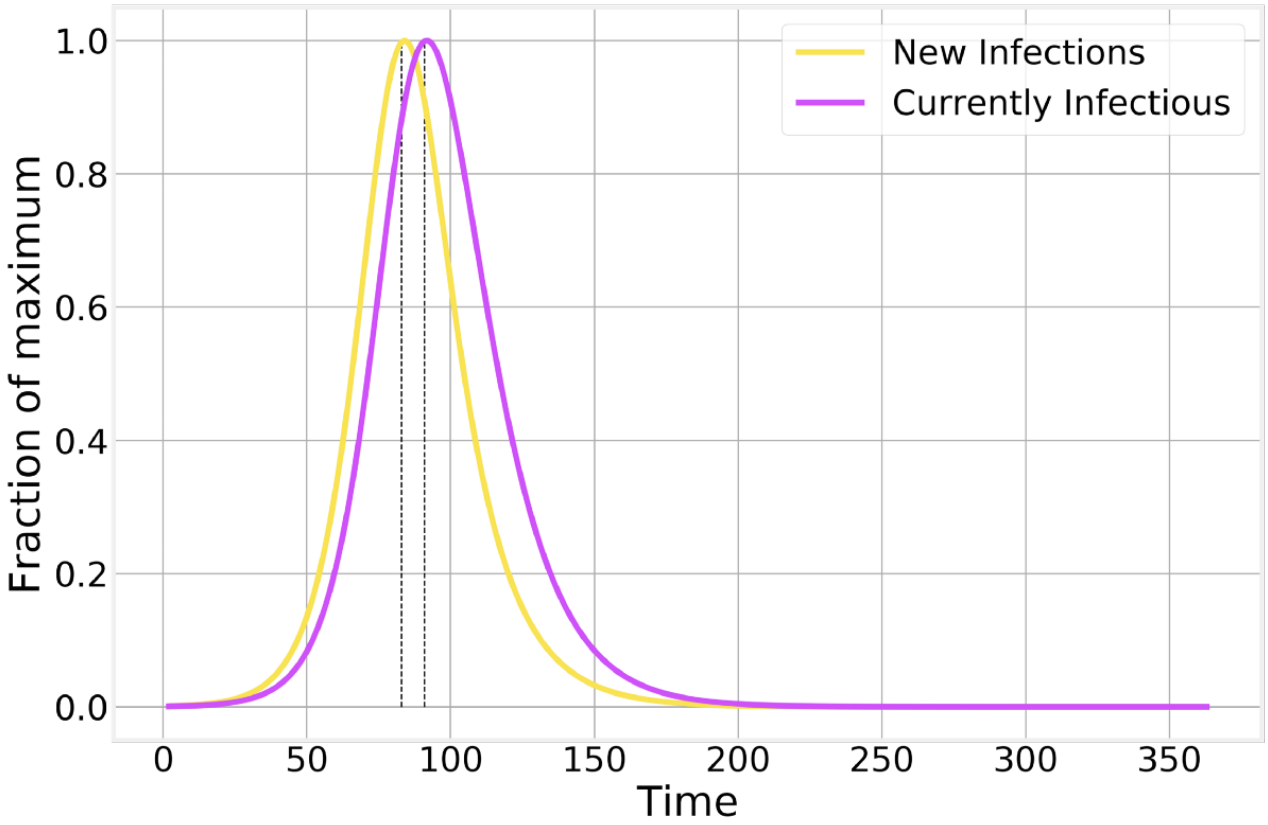
Effect of time dependent testing rate

In this figure we compare the number of real infectious cases (in purple), the result of uniform testing (dashed orange line) and dynamic testing rates (solid orange line). For clarity, we plot the different curves in a logarithmic scale (the change from one horizontal grid line to the next corresponds to a factor of 10x) and include an exponential fit line (thin blue line) as a guide to the eye that represents the overall exponential trend.

Dynamic lags

Another important factor to consider is the temporal evolution that is intrinsic to the disease progression. A healthy individual comes in contact with an infectious person and becomes infected. Her infection will last for a specific number of days, meaning that the current number of infectious individual is the sum of everyone who got infected today, yesterday, the day before, etc... and hasn't had time to recover yet.

This implies that there is a natural lag between the peak of new infections and the peak in the total number of infectious individuals that is proportional to the duration of the infectious period.



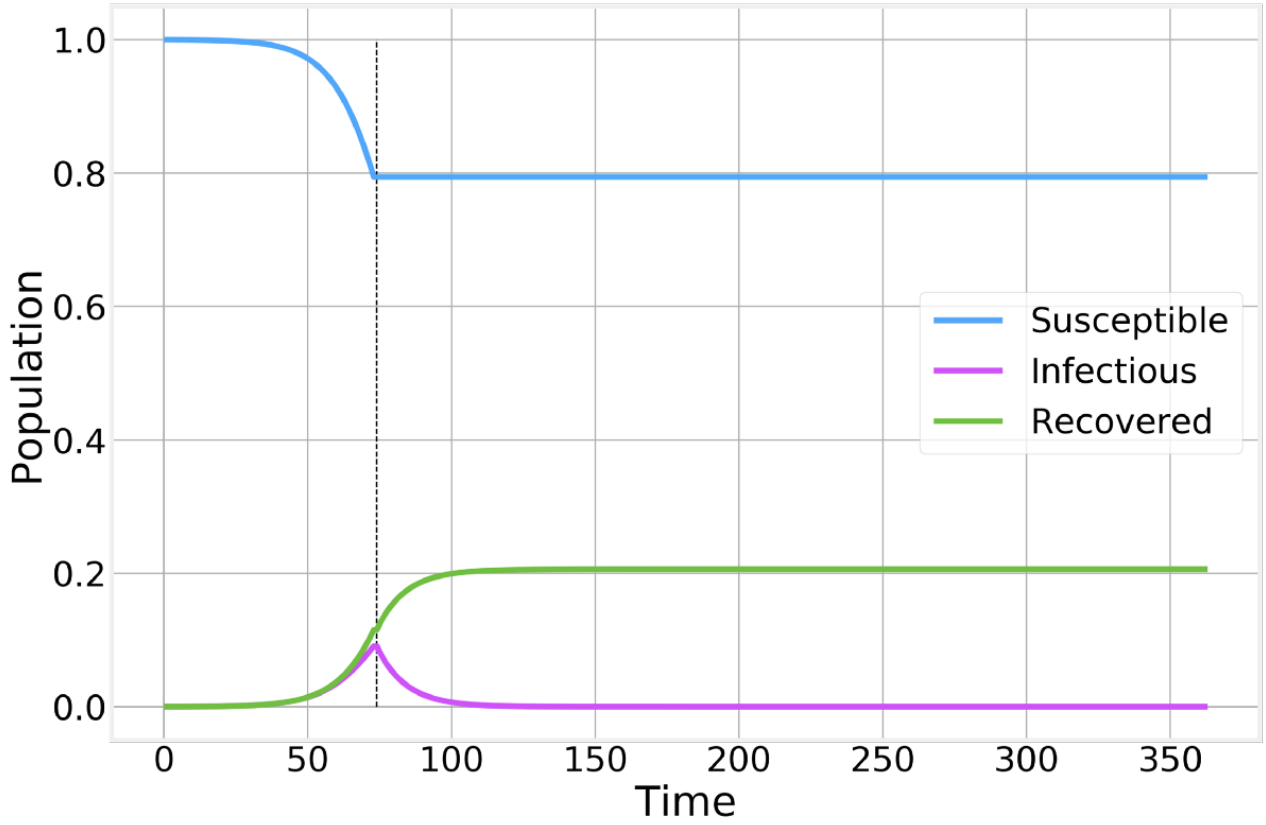
Lag between the peak in new infections and in the number of currently infections individuals

One important consequence of this lag is that even if the number of new infections today is smaller than it was yesterday and the day before, it will take several days before the effects are noticeable as a reduction in the total number of infected cases.

Lockdown procedures

As the epidemic has progressed, many countries around the world, starting with China, have tried to implement lockdown or quarantine procedures to try to contain the spread of the disease. These measures have proven unpopular with the public due to their social and economic consequences, so it is important to understand the effect they have in stopping the epidemic spreading.

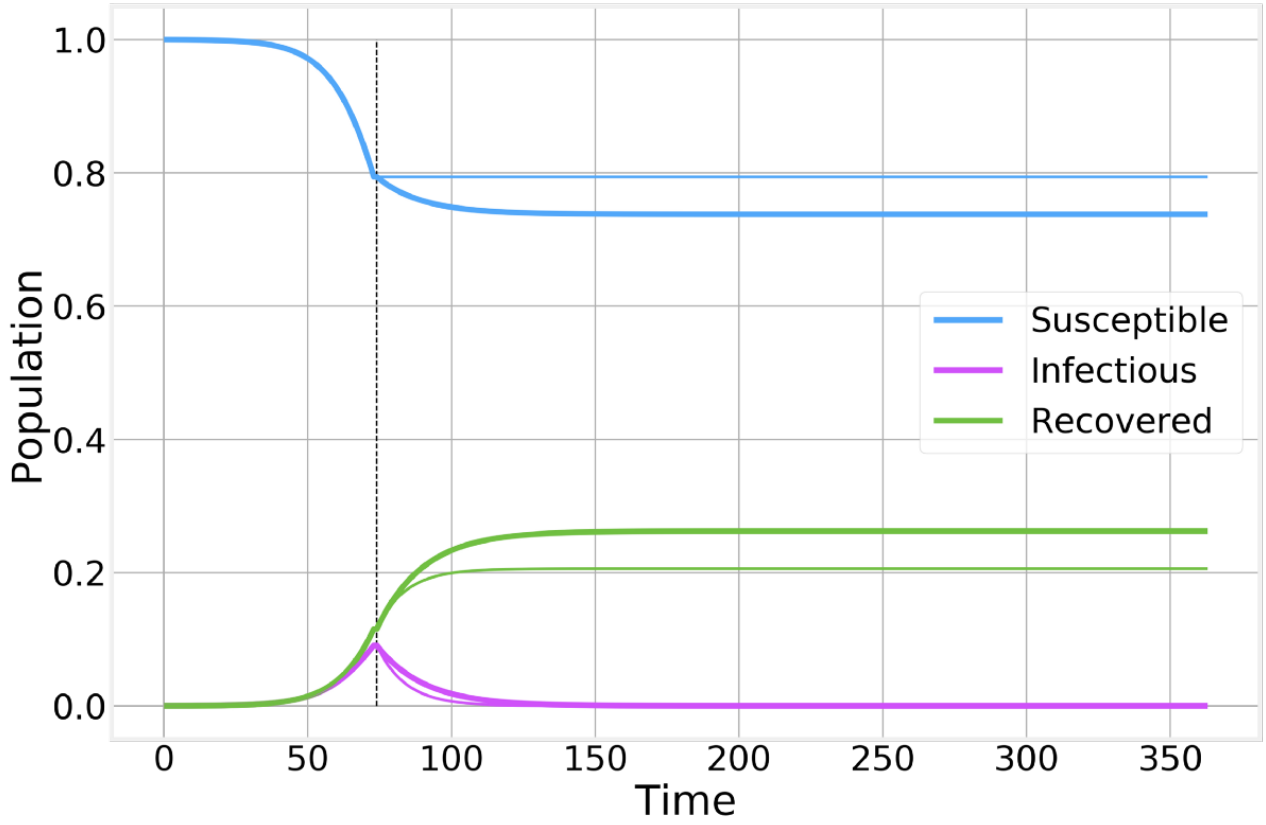
Let's imagine the perfect containment scenario. I wave a magic wand and every one stays home, exactly 6 feet away from each other at all times and no new infections can be generated. In our SIR framework, this corresponds to suddenly setting $R_0=0$ or simply eliminating the interaction transition from the model. The results are stunning:



Perfect containment strategy. Strategy is implemented at the time indicated by the vertical dashed line and maintained as long as necessary for the number of infectious individuals to reach zero.

While no new infections are generated, the total number of infected individuals still remains high for several weeks as the currently affected people gradually recover from the disease.

Naturally, no containment strategy is perfect, but let's say we do a pretty good job and instead of driving the R_o to *0* we managed to drive it to *0.5*. As we've shown above, whenever $R_o<1$ the epidemic starts to die off, but it takes significantly longer than in the ideal scenario and results in a larger number of total infections:



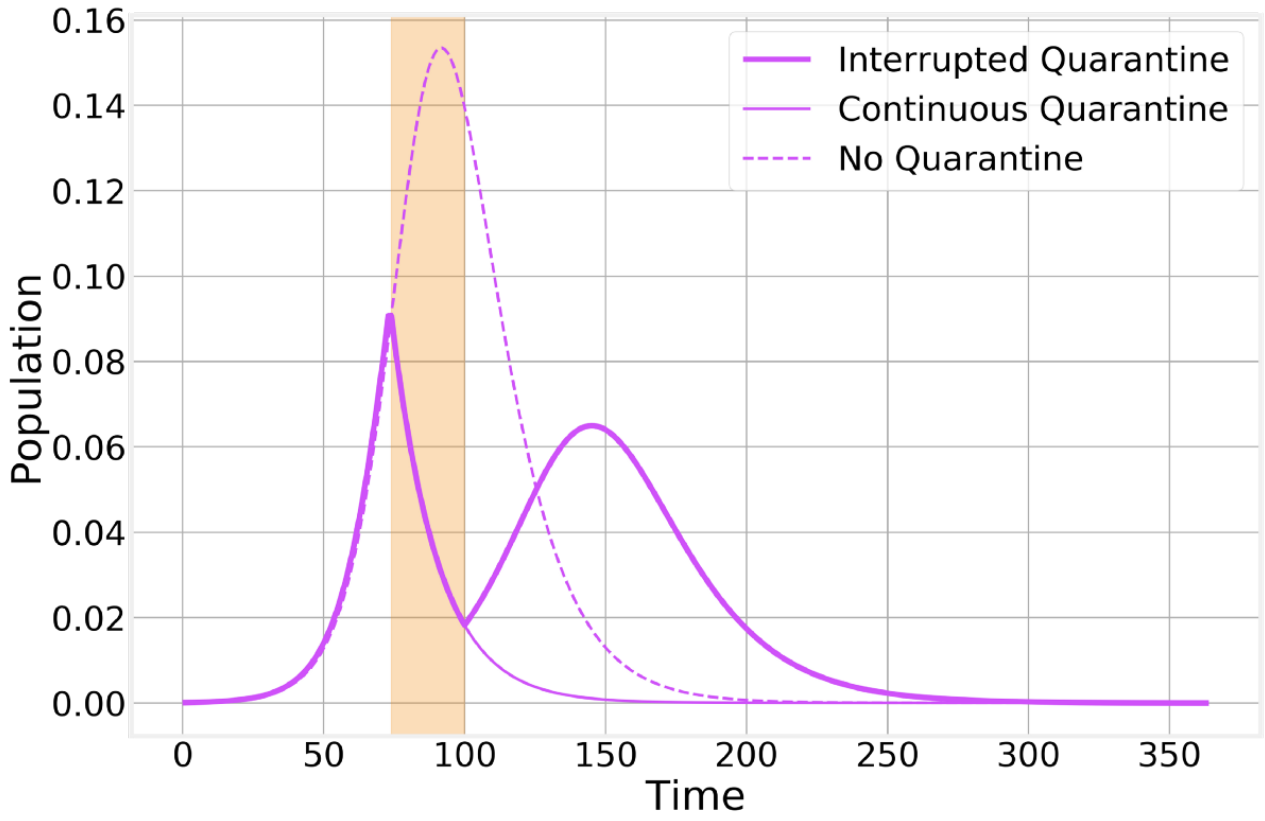
Imperfect containment strategy. Strategy is implemented at the time indicated by the vertical line and maintained for as long as necessary for the number of infected to reach zero. Thin solid lines correspond to the previous perfect scenario and are shown for comparison.

If however, for some reason, the social or economic costs of the lockdown are deemed to be too costly and the quarantine is lifted prematurely we simply return to the previous, unrestrained, epidemic spreading scenario:

Imperfect containment strategy. Strategy is implemented at the time indicated by the vertical shaded area. Dashed and thin solid lines correspond to the no-intervention and imperfect lockdown scenarios, respectively, and are shown for comparison.

As we can see, a prematurely broken lockdown quickly results in a second wave of the epidemic leading to **almost** as many total cases as if there had been no intervention whatsoever. However, it does still have the benefit of keeping the peak number of sick individuals below what would normally be and a “spreading out” of the epidemic curve: In other words, the flattening of the curve that will help **prevent the overwhelming of the healthcare system** .

For clarity, let’s also take a look at just the number of infectious cases



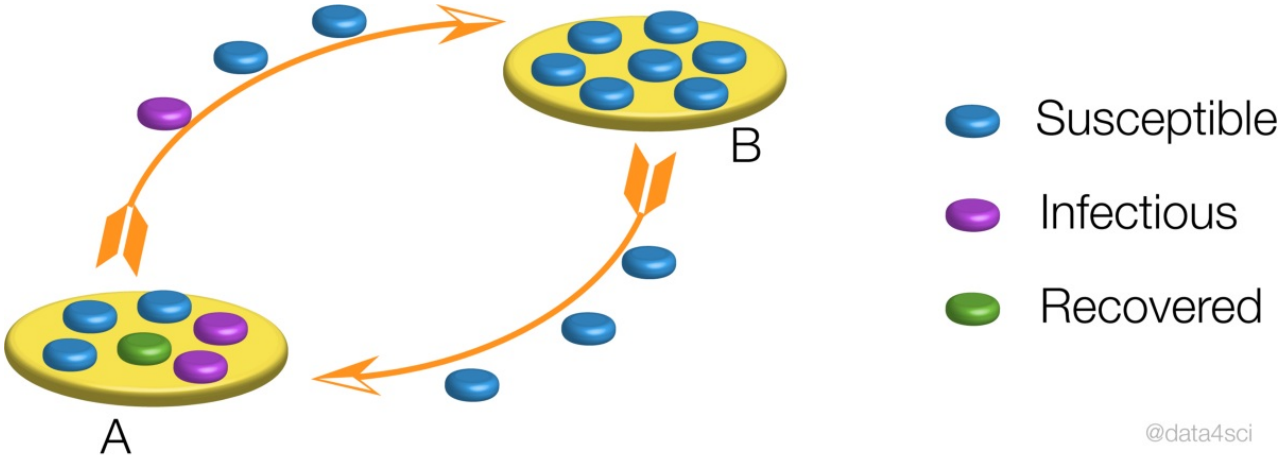
Imperfect containment strategy. Strategy is implemented at the time indicated by the vertical shaded area. Dashed and thin solid lines correspond to the no-intervention and imperfect lockdown scenarios, respectively, and are shown for comparison.

It is not for a poor Physicist such as myself to opine on whether or not the current world wide shutdown is worth it economically or socially. The best I can do is help you understand better its practical effects.

Structured populations

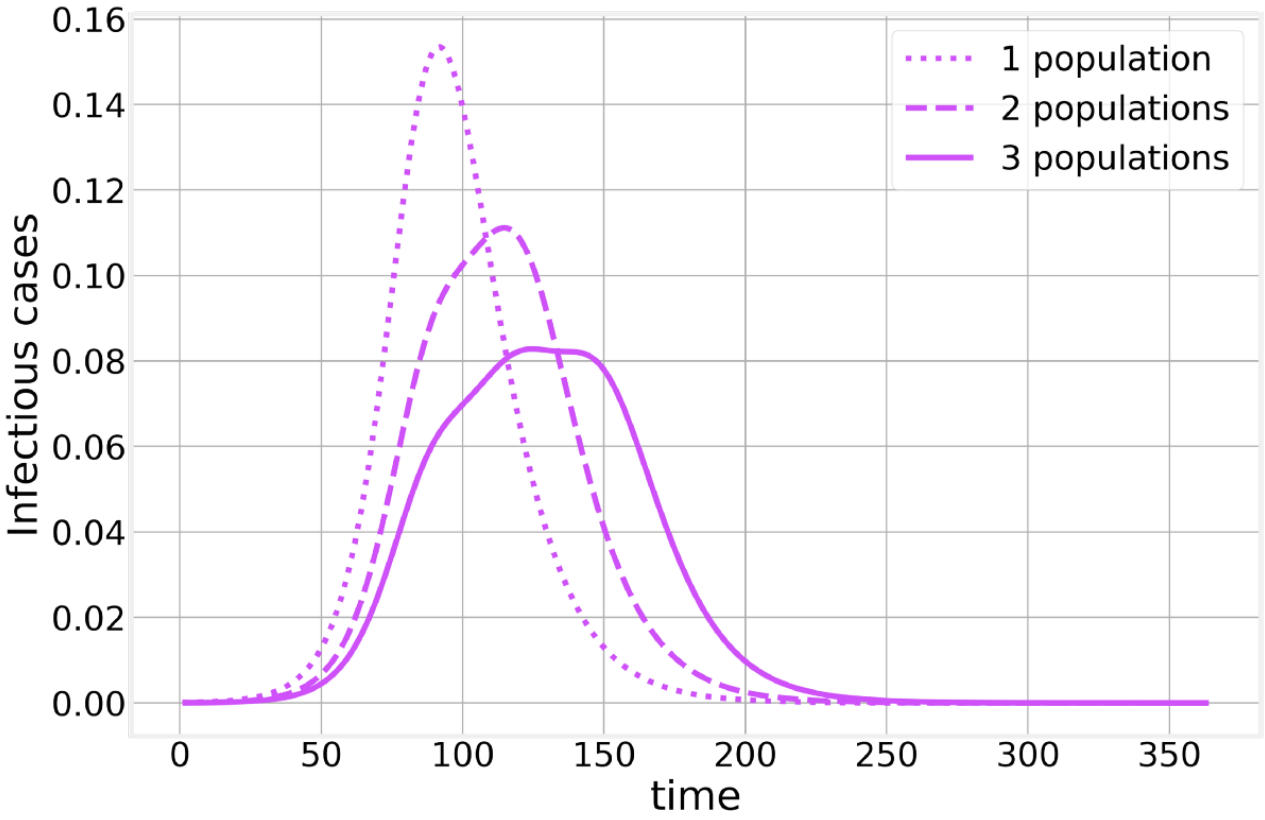
This post is already extremely long, but I would like to consider one extra point. Compartmental models, by their very nature make significant simplifications and assumptions. One fundamental assumption is that the underlying population is well mixed: every individual is in potential contact with any other individual. While this is clearly false for any large population, it is often a good enough approximation for qualitative analysis of epidemic dynamics.

However, if we try to overextend this kind of models, we quickly discover that countries and cities are not homogeneous populations. Countries are made up of states, states are constituted by cities and rural areas, etc.



Schematic representation of the epidemic in between neighboring populations.

Within each population, the epidemic will proceed as we have described above but when we combine multiple populations the results are much less clear. Let us consider two populations, say two neighboring cities. The epidemic starts in one of them and through commuting or traveling, eventually, one infectious individual will infect the neighboring city, resulting in a timing difference between the two populations. If we naively treat these multiple populations as a single one (as when looking only at state or country totals) the resulting curve is strongly affected by the timing difference between the two populations, resulting in epidemic curves that bare little to no similarity to the simple examples we've analyzed so far, making any time of exponential fitting an idle pursuit with little to no practical use.



Resources

If you've made it this far, congratulations. You now know more about epidemic modeling than most fearless curve fitters out there and hopefully you won't commit the same mistakes they're making.

And if you're still craving for more, Part II of this blog series is already published: Epidemic Modeling 102: All CoVID-19 models are wrong, but some are useful and you should check it out.

All the code necessary to implement the models described above and generate the figures used can be found in this posts GitHub repository:

DataForScience/Epidemiology101

Contribute to DataForScience/Epidemiology101 development by creating an account on GitHub.

github.com

If you enjoyed this post, you might also enjoy my weekly newsletter where I share the latest news and developments in Machine Learning and Data Science as well as any future blog posts I write.

Discover Medium

Welcome to a place where words matter. On Medium, smart voices and original ideas take center stage - with no ads in sight. Watch

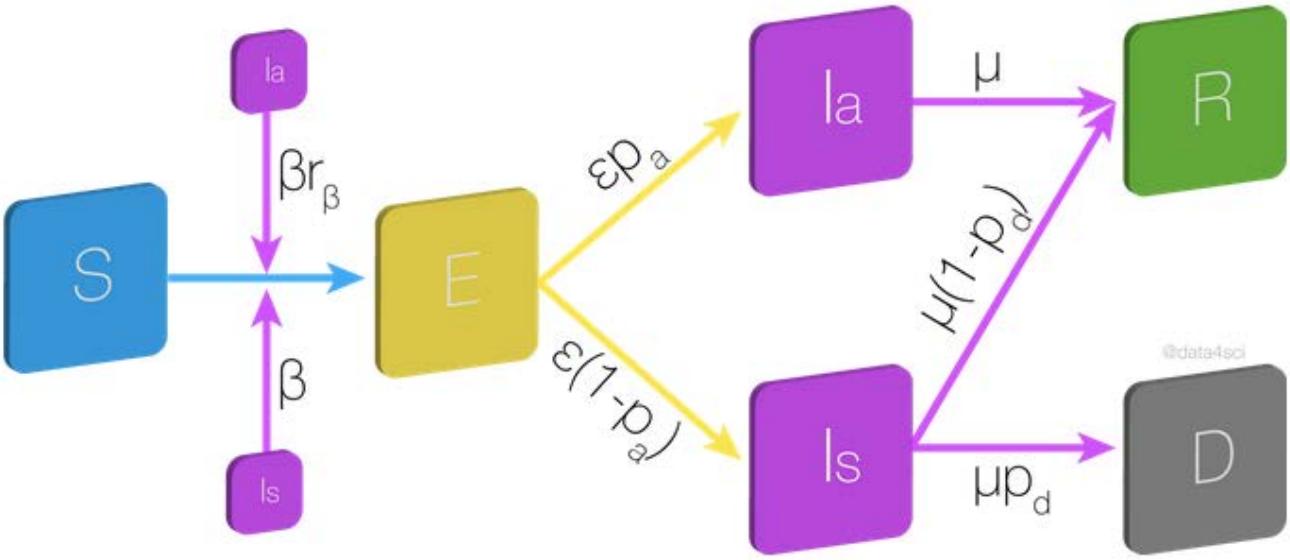
Make Medium yours

Follow all the topics you care about, and we'll deliver the best stories for you to your homepage and inbox. Explore

Become a member

Get unlimited access to the best stories on Medium — and support writers while you're at it. Just $5/month. Upgrade

2. All CoVID-19 models are wrong, but some are useful



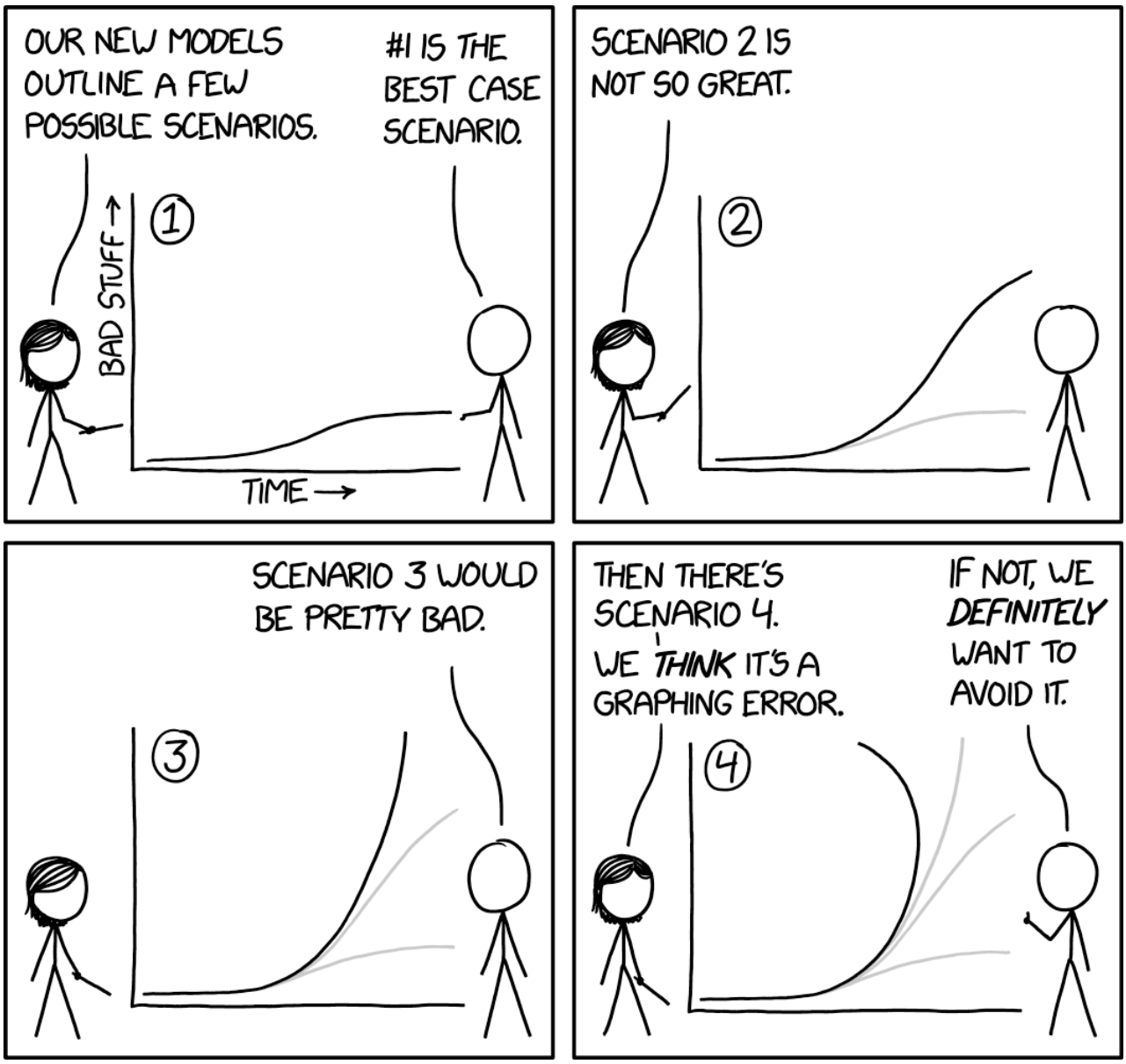
What follows is my personal perspective, as an individual with some real world experience in epidemic modeling during previous pandemics and shouldn't reflect on any group or institution I might be affiliated with.

So, without further ado...

Models vs the real world

As George E. P. Box, a statistician, famously said "all models are wrong, but some are useful". This is perhaps never more true than during a crisis. Information is limited, often wrong, but decisions must be made and implemented based on what is known at the time.

It is also during an ongoing crisis that models play their most fundamental role, that of allowing us to explore scenarios and work through the consequences of our decisions:

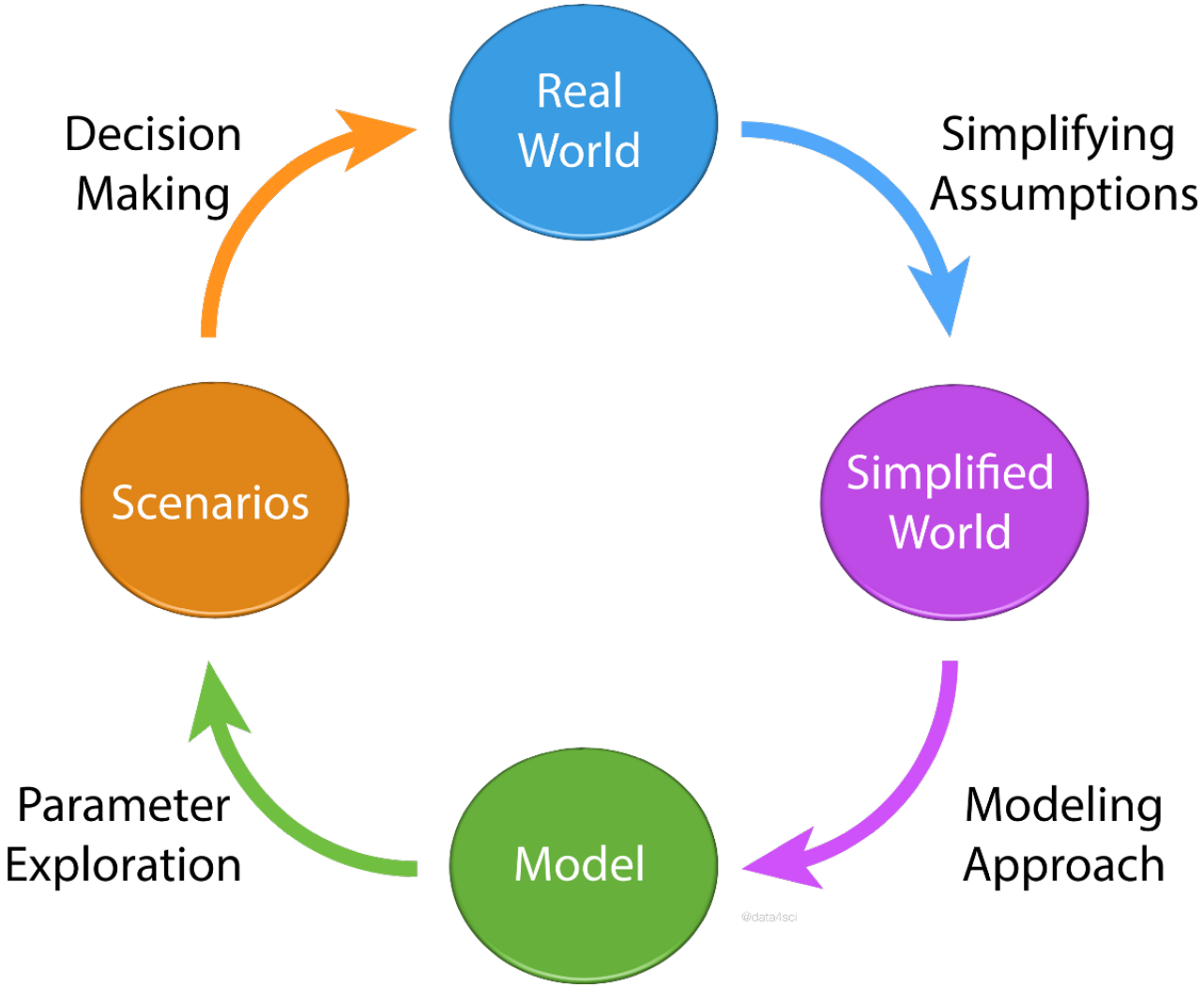


XKCD: "Remember, models aren't for telling you facts, they're for exploring dynamics. This model apparently explores time travel"

However, care must be taken to avoid mistaking the model for the reality. After all, "the map is not the territory". The development of a model, regardless of the domain of application, typically follows a common pattern:

A simplified version of the world is created, to which a specific modeling approach can be applied, resulting in a working model. The simplifications made can be due to a variety of factors such as the lack of specific data, excessive complexity, intractability, among others. The modeling approach chosen is both influenced by and helps drive the assumptions that are made, often resulting in the stereotypical overabundance of Physicists concerned about Spherical Cows or trying to apply Ising Spins to every possible problem.

Once a working model is obtained, we can use it to explore scenarios, the consequences of specific decisions, etc. Finally, it is by studying the scenarios that are outputted by our models that decisions are taken in the Real World. Graphically, we have:



Naturally, this is a simplified and schematic view (and a model in and of itself) to help illustrate the various points at which our modeling efforts can go awry, leading the results of our models to differ from what we actually observe in the real world.

While in many cases, mismatches between the model and reality can be traced back to errors made during the process, they can also be due to the fact that our model was successful and it resulted in appropriate measures being taken to prevent the undesirable scenarios it predicted. This is specially true in the case of highly visible models that are used to guide government decisions and interventions such as in the case of an ongoing pandemic like the one we're living through now:

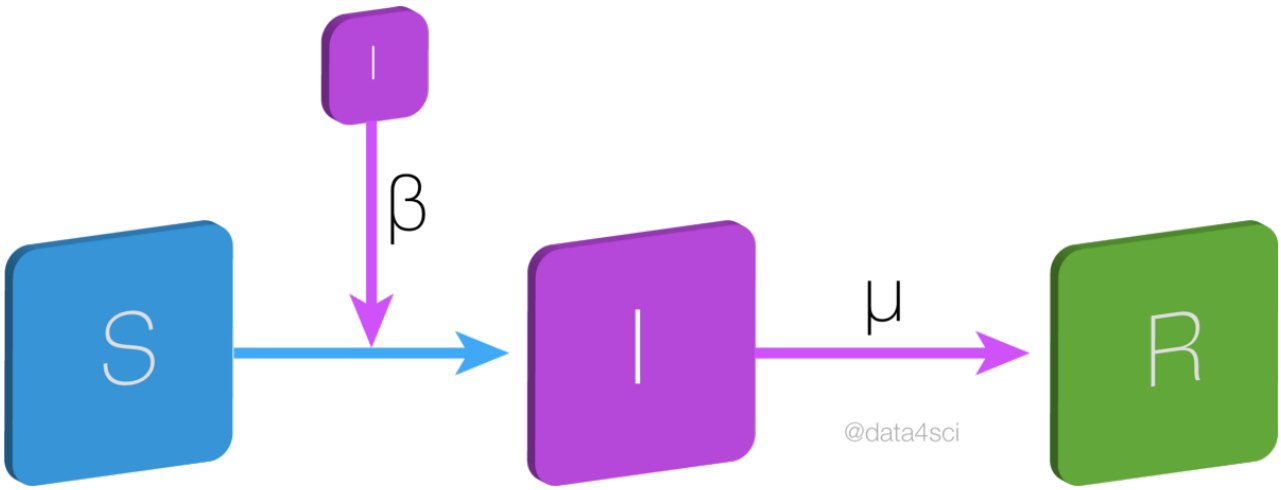
> "The most important function of epidemiological models is as a simulation, a way to see our potential futures ahead of time, and how that interacts with the choices we make today. With COVID-19 models, we have one simple, urgent goal: to ignore all the optimistic branches and that thick trunk in the middle representing the most likely outcomes. Instead, we need to focus on the branches representing the worst outcomes, and prune them with all our might. Social isolation reduces transmission, and slows the spread of the disease. In doing so, it chops off branches that represent some of the worst futures. Contact tracing catches people before they infect others, pruning more branches that represent unchecked catastrophes." — Zeynep Tufecki, The Atlantic

It is this kind of misunderstanding that leads to public distrust in scientific models in particular and Science in general.

My hope is that this (and many other posts out there) can help the general public to understand the underlying assumptions, power, and limitations of scientific models and how they can be put to good use.

Susceptible-Infectious-Recovered (SIR) Model

Now that we have established both the advantages and limitations of using models to understand the world, we can start exploring how to improve the simple models we introduced in the previous post.



SIR Model

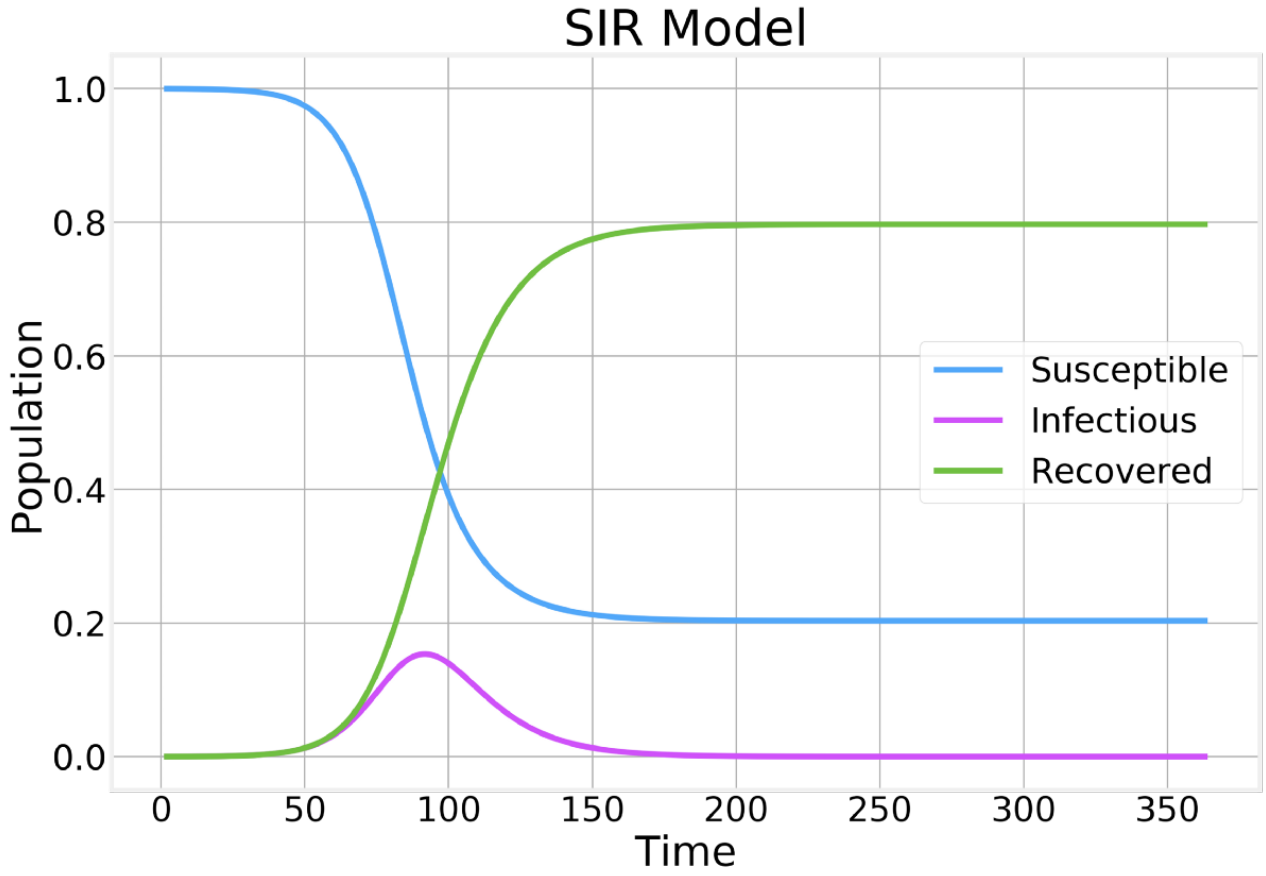
The SIR model is one of the simplest and best known epidemic models. Its popularity is due, in no small part, to its ability to establish a perfect balance between simplicity and usefulness. It is still relatively amenable to mathematical and analytical exploration while at the same time it is able to capture the fundamental features of the epidemic process: healthy (*Susceptible*) people become infected when coming in contact with *Infectious* individuals only to eventually *Recover* after a certain period of time. The process is illustrated schematically in the figure at the top of this section.

This model can be written mathematically using a simple set of partial differential equations:

Which can be numerically integrated to obtain the values of each compartment as a function of time, just as done in the previous blog post:

$$\begin{aligned}
\frac{\partial}{\partial t} S_t &= -\beta S_t \frac{I_t}{N} \\
\frac{\partial}{\partial t} I_t &= +\beta S_t \frac{I_t}{N} - \mu I_t \\
\frac{\partial}{\partial t} R_t &= +\mu I_t
\end{aligned}$$

Susceptible-Infectious-Recovered model



Fraction of the population in each compartment as a function of time

While this kind of equations can be useful to explore analytical results for simple models like the SIR model, they quickly become unwieldy for more complex models. However, it is easy to note how they have a one-to-one correspondence with the illustration above:

> Interactions correspond to terms involving two compartments and the total number of individuals in the population:

$$\beta S_t \frac{I_t}{N}$$

Interaction term

> While spontaneous transitions correspond to terms involving just a single compartment:

$$\mu I$$

Spontaneous term

> The sign of each term is determined by whether the equation we are considering corresponds to the "source" or "target" compartments. Notably, "agent" compartments are not affected unless they are also "targets".

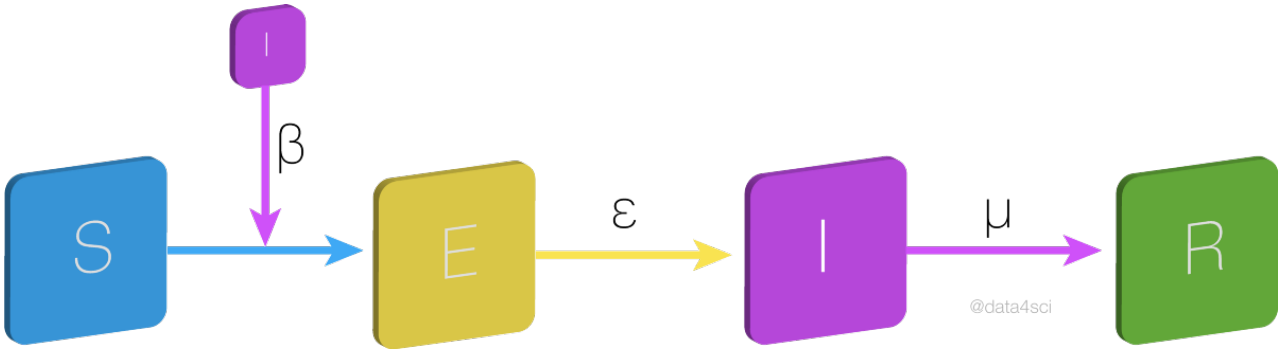
This one-to-one correspondence between transitions and terms allows us to simply "draw up" arbitrarily complex models that can be trivially implemented using generic code (like the one in EpiModel.py) without having to write out and debug all the rules "by hand".

In the rest of the discussion we will focus on discussing the assumptions and details of the various models while avoiding as much as possible the use of complex mathematical expressions.

Incubation Period

One of the main limitations of the SIR model is the fact that the infection develops instantaneously without any incubation period what so ever. You'll recall from recent news that this is not a very realistic scenario and the incubation or latent period is one of the most important factors that must be understood in order to contain an epidemic: For how long must a suspected case be kept under watch until we can be certain that the person will not become infectious?

We can address this limitation by adding one extra step (compartment) to our epidemic model: The *Exposed* (or *Latent*) compartment. When a *Susceptible* person comes in contact with an infectious one s/he moves to the *Exposed* from which s/he transitions to the *Infectious* compartment at a fixed rate ε. While in the *Exposed* compartment the person is said to be "incubating" the disease, possibly even starting to develop symptoms, but is not yet able to infect other individuals. The resulting model is known as the *Susceptible-Exposed-Infectious-Recovered* (SEIR) model:



SEIR Model

Here we have 4 distinct compartments connected by one interacting transition and two spontaneous ones:

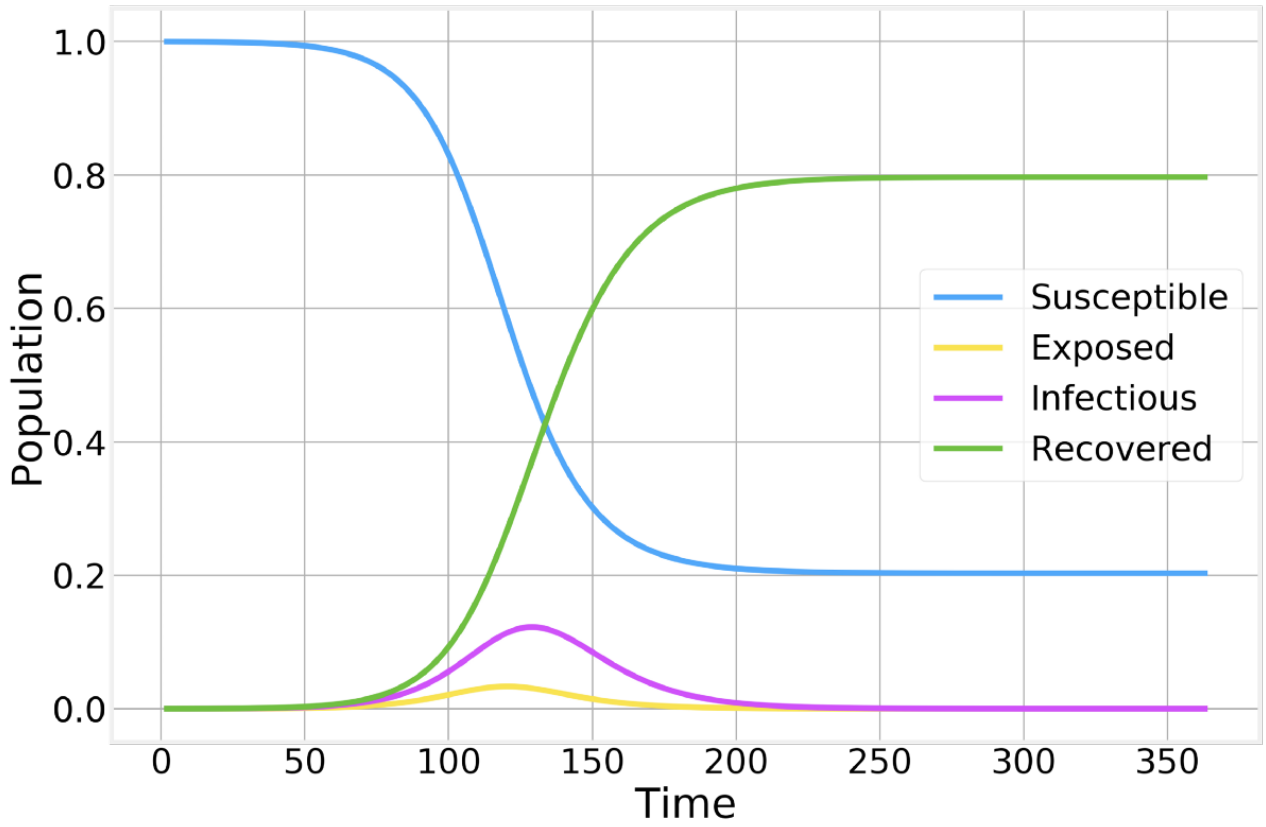
And the evolution of the various compartments is simply:

$$S + I \xrightarrow{\beta} E$$

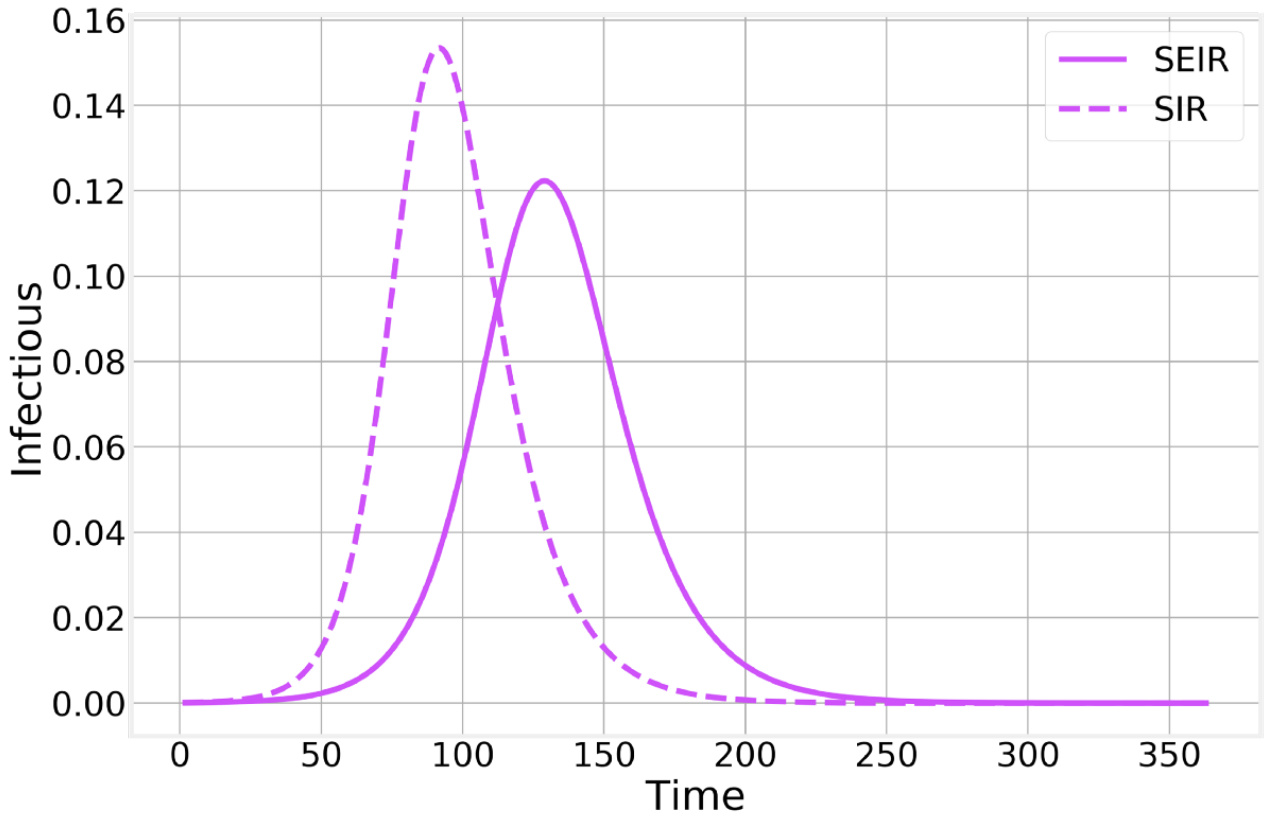
$$E \xrightarrow{\epsilon} I$$

$$I \xrightarrow{\mu} R$$

SEIR Transitions



Here we highlight that the addition of the extra compartment didn't change the total number of people who become infected, but it does have a strong impact on the temporal evolution of the epidemic, significantly delaying and widening the peak of infectious cases. It effectively "flattens" the curve:



Epidemic peak comparison between the SIR and SEIR models.

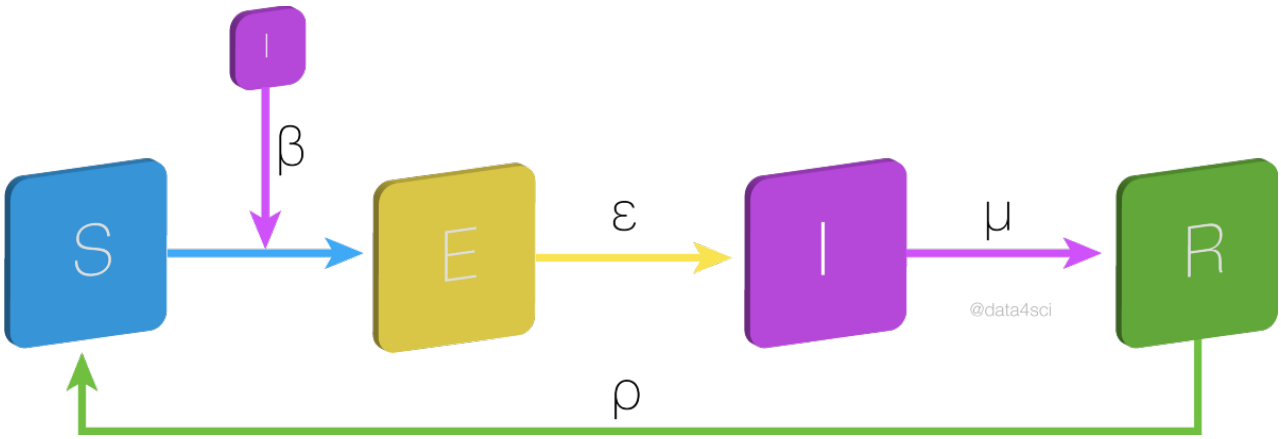
It should be clear how this has a direct impact on the likelihood of the healthcare system being overwhelmed and the necessary duration of any quarantine measures imposed: **a lower peak reduces the stress in the healthcare system**, while **a longer duration implies that longer period of social distancing** is necessary.

Temporary Immunity

Another fundamental assumption underlying the SIR model is the idea that *Recovered* persons are permanently immune from the disease. While this is the case with many common diseases, there have been some reports of CoVID-19 patients being re-infected after recovery.

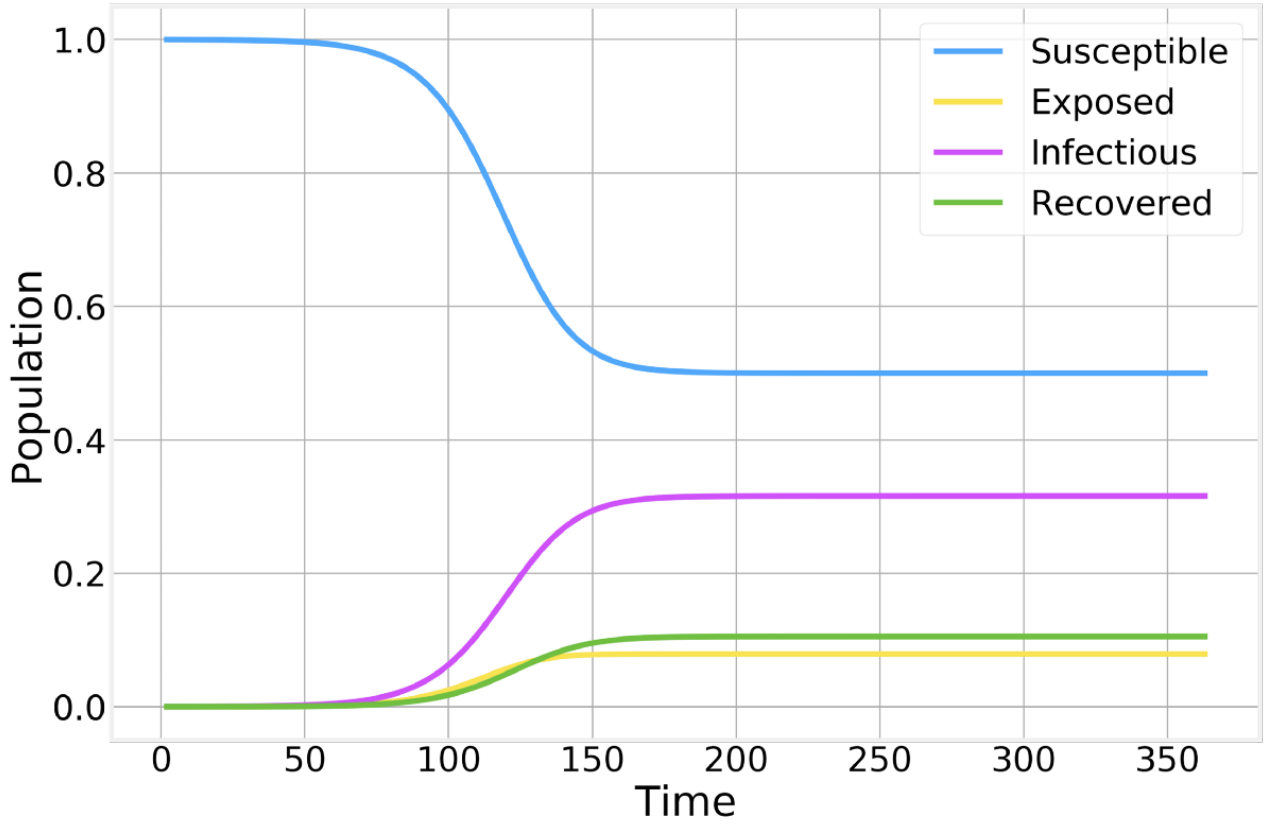
Re-infection in such a short period of time is unlikely (even temporary immunity typically lasts for a few months or years) and these cases are more likely to be due to faulty tests, but it is certainly a possibility that should be considered.

By simply adding a spontaneous transition from the *Recovered* compartment back to the *Susceptible* compartment, we obtain the **SEIRS** (can you guess what the letters stand for? 😀):



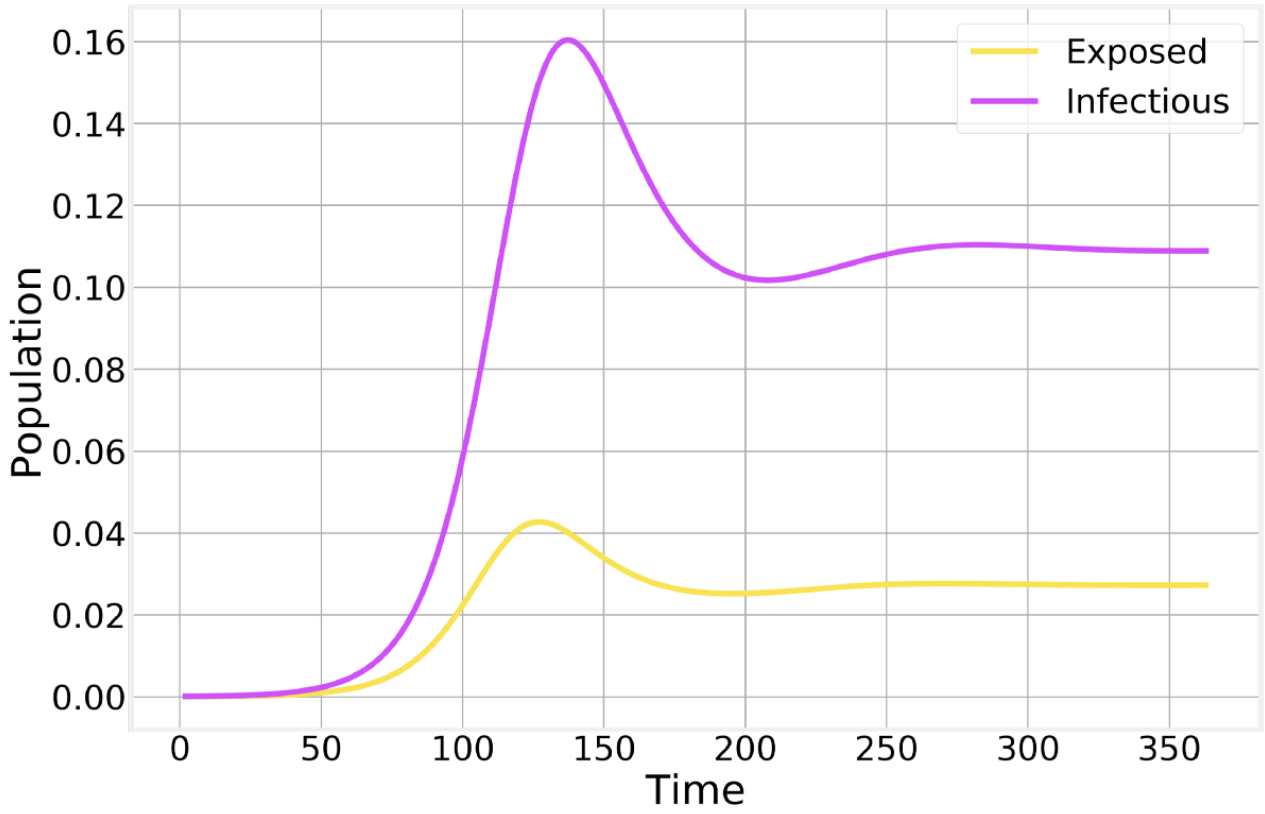
SEIRS Model

This seemingly innocuous addition to the model has a **very important effect**. By allowing *Recovered* individuals to once more become *Susceptible*, we replenish the group of people that can once again be infected. The end result is that the epidemic never burns itself out (its fuel is never exhausted) and the disease becomes endemic, with a constant fraction of the population remaining infected!



Endemic final state of the SEIRS model

The rate ρ at which immunity is lost has a determinant effect in the progress of the epidemic and the rise of endemicity. If ρ is sufficiently small (immunity is longer lasting) we can even have several epidemic peaks before the steady state of a fixed fraction of the population is reached.



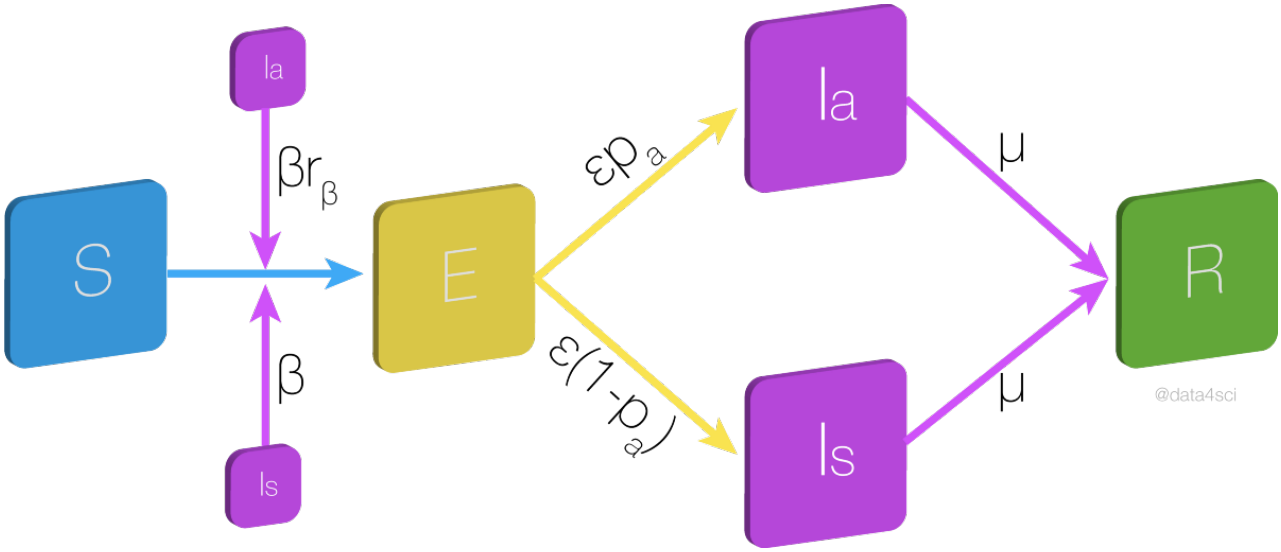
Exposed and Infectious population in the SEIRS model

The appearance of the peak is due to the fact that the temporary immunity afforded by the disease is sufficiently long to allow the epidemic to follow most of its course before the number of *Susceptibles* starts to increase again, adding fuel to the fire.

Asymptomatic Cases

In many diseases, a significant fraction of infected individuals remain asymptomatic throughout the course of the disease. In the case of seasonal Influenza, this number is typically around 33%, while for CoVID19 the number is thought to be 40% or higher, thus significantly skewing the total number of cases.

Asymptomatic individuals are often less infectious than those displaying symptoms by some fraction r_β. We can model their effect by splitting the *Infectious* compartment in two: a Symptomatic, *Is*, and an Asymptomatic, *Ia*. A fraction p_a of all of those *Exposed* become asymptomatic while the remaining $(1-p_a)$ develop symptoms. Our model is then:



As we now have two *Infectious* compartments we must also redo our R_o calculation. Fortunately, the modification is simple: since we have split the original *Infectious* compartment in two, our value of β is simply the weighted average of the original and the reduced β.

We can easily verify that if r_β is 1 we recover the original SIR value, while if r_β is 0 (the asymptomatic and completely non-infectious) we reduce the original R_o by a factor of (1-p_a) as we effectively have that much smaller *Infectious* population.

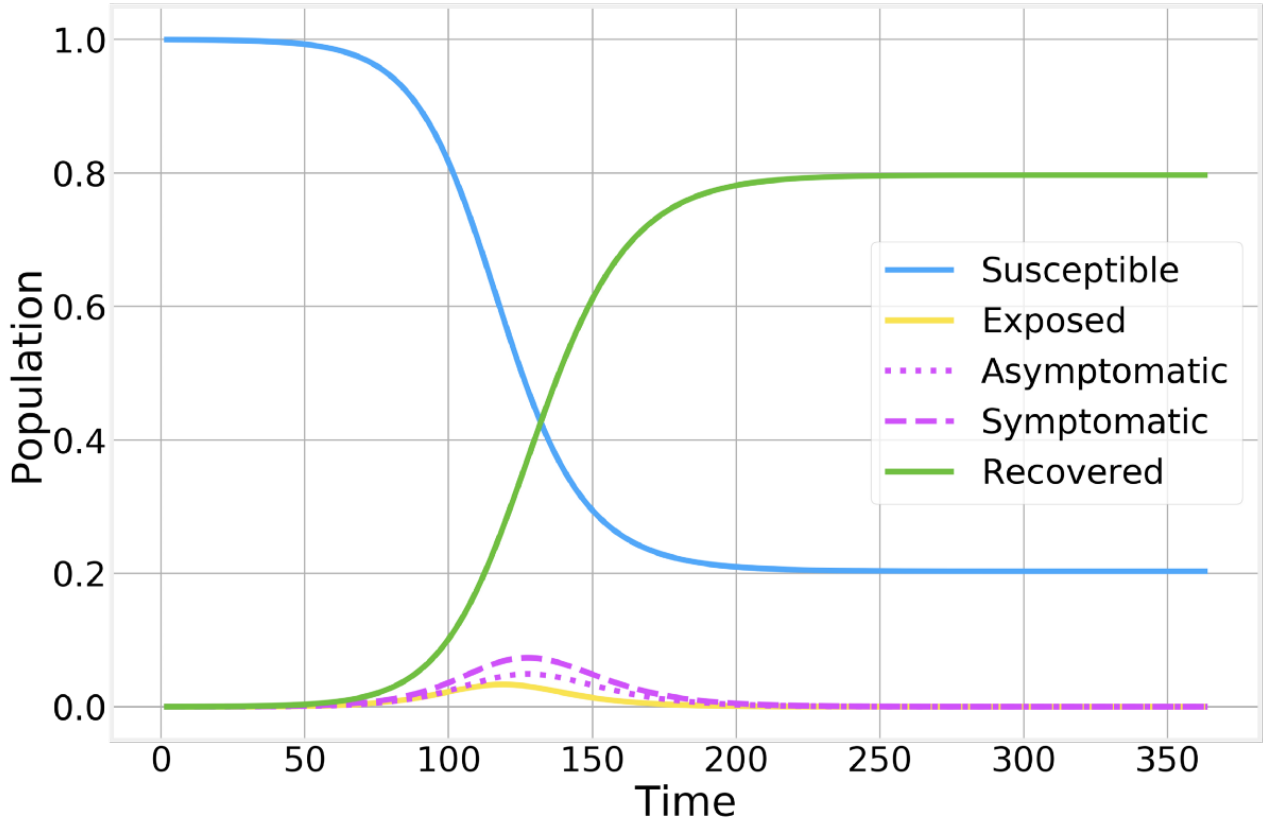
$$R_0 = \frac{\beta}{\mu}\left[p_a r_\beta + (1 - p_a)\right]$$

In order to maintain the same value of R_o we simply calculate the value of β as:

This approach makes it easier to compare the results from both models since they both have the same value of R_o.

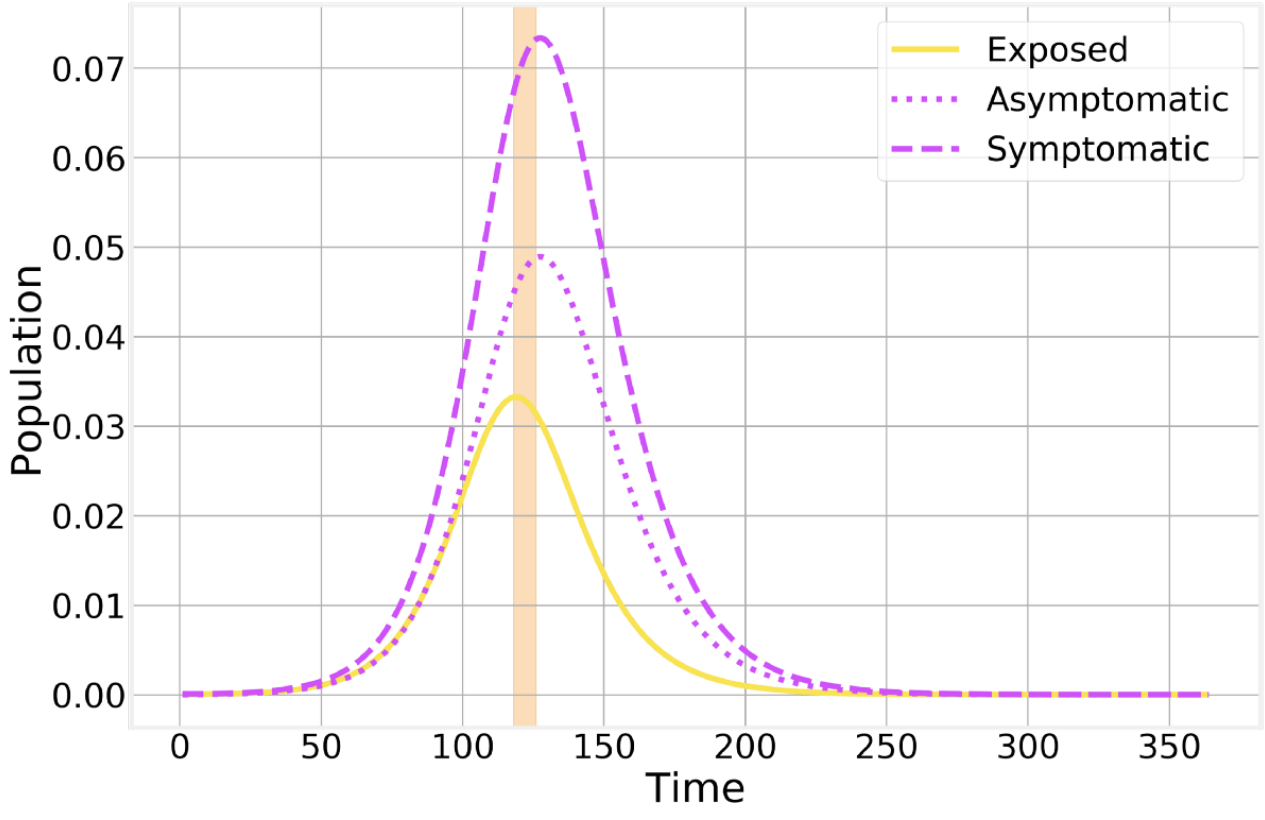
$$\beta = \frac{R_0 \mu}{p_a r_\beta + (1 - p_a)}$$

As we add more and more compartments to our models, the smaller the population of each individual compartment becomes.



Compartmental structure of the Symptomatic/Asymptomatic model

We can easily verify that the value of R_0 remains the same as before by looking at the *Recovered* and *Susceptible* curves at the end of the epidemic. On the other hand, we now have 3 distinct infected compartments, 2 of which are *Infectious* and peak at the same time and a few days after the *Exposed* population:



Peak comparison between the three infected compartments

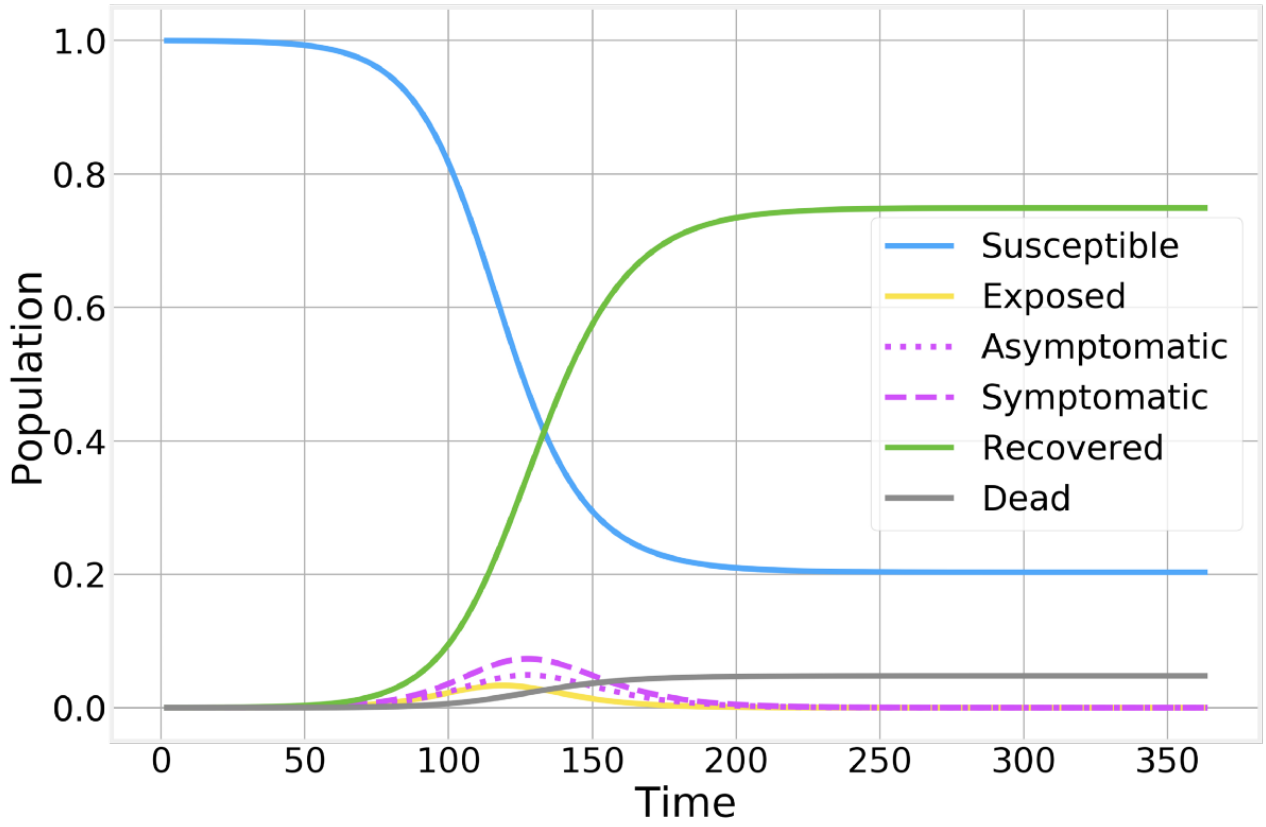
Here we should note that we explicitly decided to keep the recovery rate, μ for both Symptomatic and Asymptomatic individuals. Had we chosen them to be different then the peaks would occur at different times and the expression for R_o would have to be revised even further.

Mortality rate

Finally, we look at the effect of explicitly considering mortality. We assume that only symptomatic cases die from the disease or, similarly, that any asymptomatic individuals that do die from the disease are not counted as such. If we assume that a fraction *pd* of symptomatic cases end up dying, our model becomes:

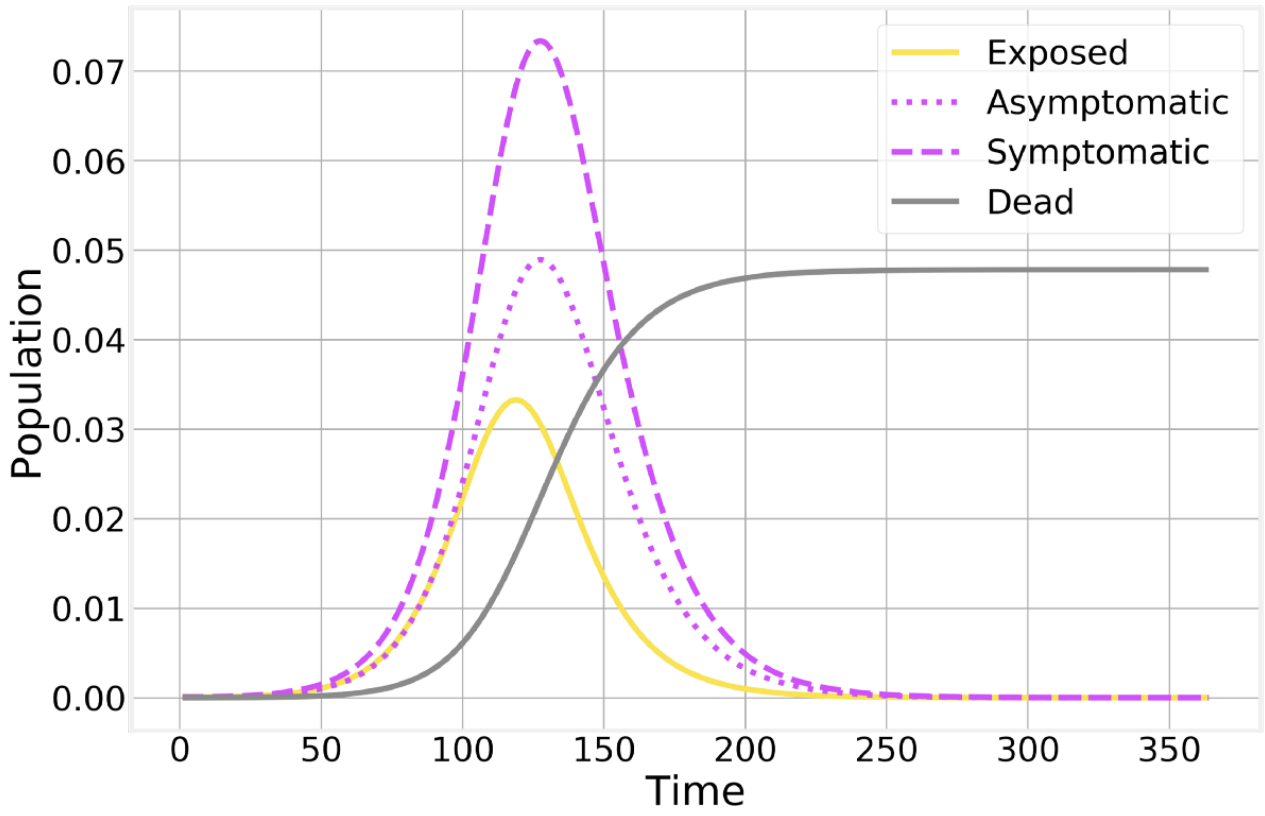
So we now have 6 compartments and a total of 7 transitions and 6 parameters, denoting how the more details we include the more complex the model becomes and the more parameters must be specified. In the early days of an epidemic most, if not all, of these parameters are partially or completely unknown. As the epidemic progresses, more and more information is gathered and more detailed models can be used. This constant refinement also helps improve the reliability of the scenarios we are able to analyze and the decisions made.

If we assume that *10%* of the symptomatic cases eventually die, we have:



It should be noted that ***10%* mortality rate is huge and unrealistic** for the kind of diseases we are considering. The reason we choose such a large number is **to make the effects obvious when plotting**.

By including the possibility of *Death*, the number of *Recovered* individuals is naturally reduced, despite the fact that none of the disease parameters have been changed. If we focus on just the relationship between the most significant compartments we have:



The total number of dead can be easily estimated. We know that for our set of parameters, 80% of the population eventually becomes infected. Of those, 60% are symptomatic and of those, 10% eventually die, so we expect that the total number of fatal cases to be 4.8% as shown in the plot above.

This value is significantly smaller that the actual mortality rate for the symptomatic cases. This is due to the fact that the number of recovered is inflated by the milder asymptomatic cases.