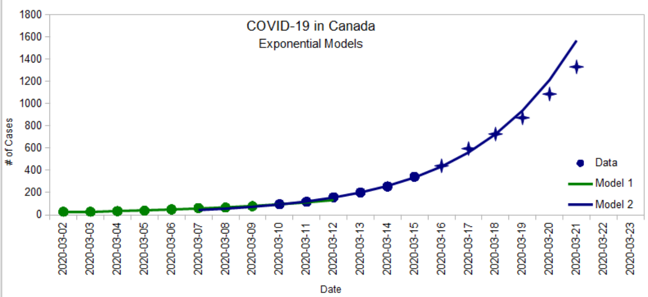
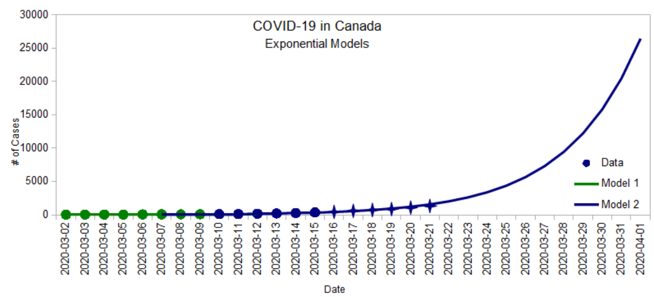
COVID-19 Spread (Part II)

- I'm not an epidemiologist, doctor, or any kind of expert on the subject. I just look at the numbers.

In Part I I built an exponential model using data between March 2 and March 15, then continued to add daily numbers to see how that model tracked:





The formulae for the exponential curves are:

- $N = 24.5\times2^{(\frac{t}

{4.1})}$ for the green line (where *t* is the number of days since

March 2)

- $N = 93.1 \times 2^{(\frac{t}{2.7})}$ for the bluel ine (where *t* is the number of days since March 1)

Initially, the number of cases doubled every 2.7 days, predicting almost 1600 cases by the end of Saturday, but since Thursday, the rate of infection seems to have slowed down a bit, which is what I'm exploring here.

Growth Factor

There's a ratio involving three data points that's useful to track how “fast” the exponential grows. It's easier to explain with an example so suppose we had three days like this:

- If the growth factor is above 1, it means we're still in the exponential phase, growing faster and faster day after day.
- If the growth factor is 1, then the infection is growing at a constant rate. This is the middle of the Logistic Curve (more on that soon).
- If the growth factor is less than 1, then the infection rate is levelling off.

To calculate the growth factor:

- The first step is to calculate the number of new cases from one day to the next.
- Then, we take the ratio between the number of new cases (20 ÷ 10 = 2)

Day	# of Cases	New Cases	Growth Factor
Day1	100		
Day2	110	10	
Day3	130	20	2

Here's the number of cases in Canada with the calculated growth factor:

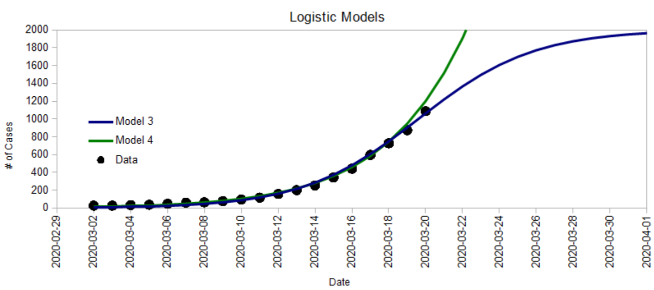
Date	# of Cases	New Cases	Growth Factor
2020-03-01	?		
2020-03-02	27		
2020-03-03	27	0	
2020-03-04	33	6	
2020-03-05	37	4	0.67
2020-03-06	48	11	2.75
2020-03-07	60	12	1.09
2020-03-08	64	4	0.33
2020-03-09	77	14	3.25
2020-03-10	95	18	1.38
2020-03-11	117	22	1.22
2020-03-12	157	40	1.82
2020-03-13	201	44	1.10
2020-03-14	254	53	1.20
2020-03-15	342	88	1.66
2020-03-16	441	99	1.33
2020-03-17	596	155	1.57
2020-03-18	727	131	0.85
2020-03-19	873	146	1.11
2020-03-20	1087	214	1.47
2020-03-21	1331	244	1.14

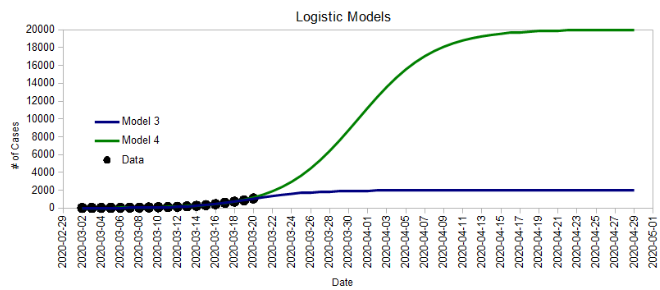
There's a lot of variation in the growth factor because real life is messy, but if we plot it on a graph, we can see a bit of a pattern:

Although the growth factor is still above 1, it looks like we might be on track to reach 1 by the end of the month. If that's the case, and if we continue to implement measures to slow the down the spread, then we'll be in a position to estimate the final outcome.

The Logistic Curve

In Part I, I showed very different Logistic Curves and explained that there's really no way of knowing which we'll follow yet. Here they are again:

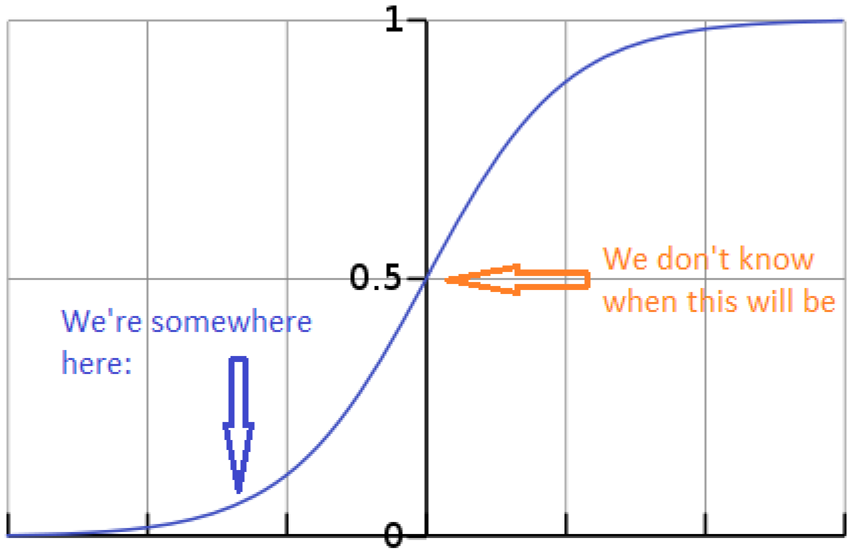




COVID-19 Spread (Part I)

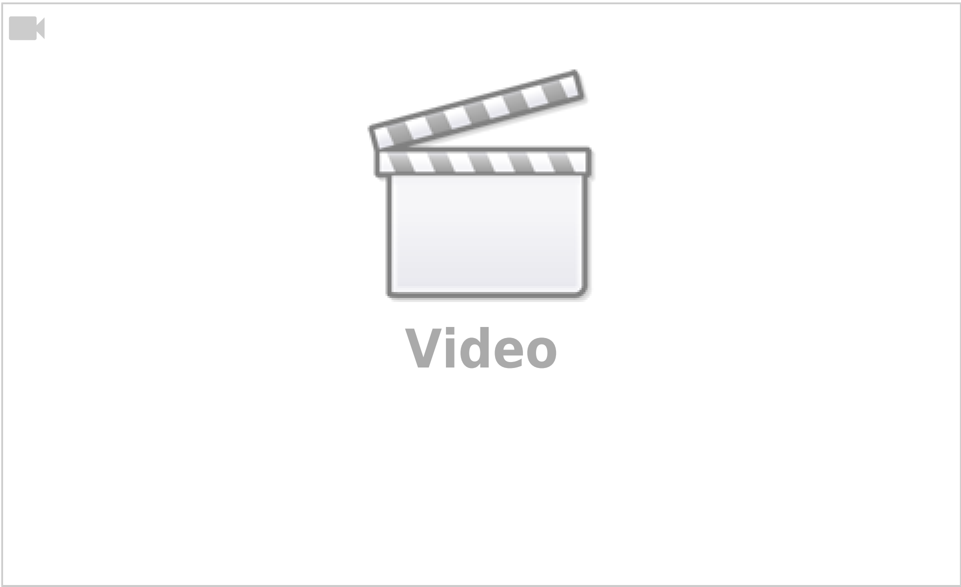
So there's a very real sense in which, *if we don't do anything different*, we could simply be about 15 days behind Italy…

But doing the right things can change that future. In reality, the spread of the infection follows more of a Logistic Function. At the beginning, it looks like an exponential, but then it flattens out. This is what the news keeps referring to when they say that social distancing and proper hand washing can help “flattening the curve” more quickly.

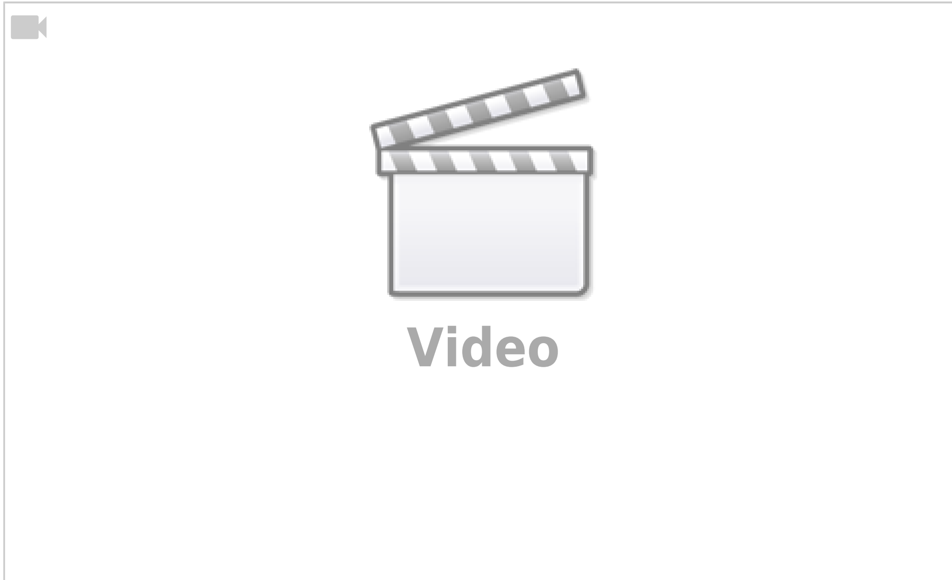


The real question is how soon will we reach that middle point, and at what height.

Here's a good video that explains this sort of math and why being able to think in exponential term is important for non-linear systems such as this one.



And here's another one with different animations that complements it very nicely.

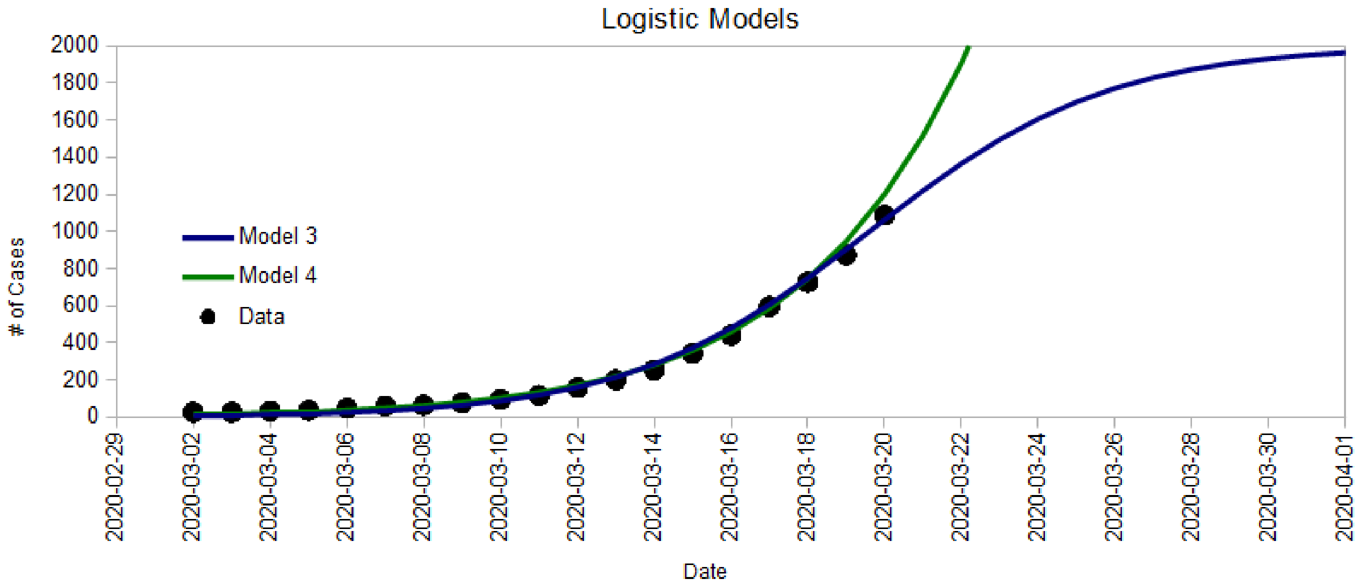


Here's an interesting article from The Washington Post showing basic random simulations for four different cases (free-for-all, attempted quarantine, mild moderate distancing, extensive social distancing).

More on the Logistic Function

This is an update from March 19th.

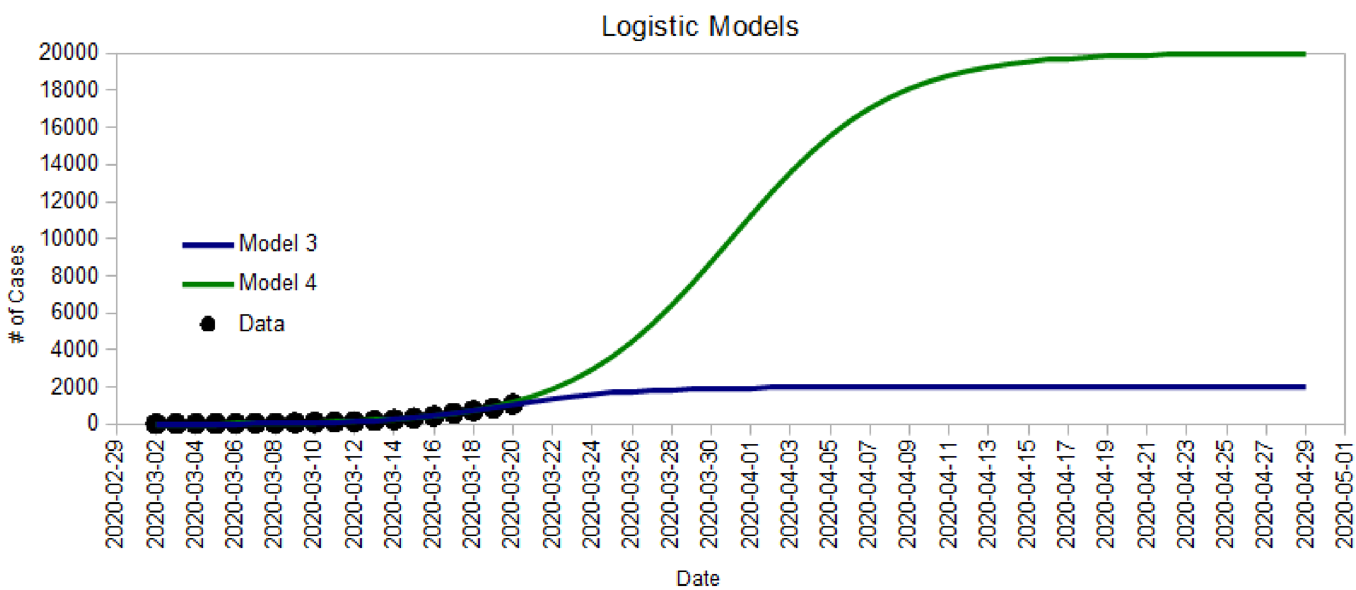
This section illustrates how eventhough the infection follows a Logistic Function, that fact alone doesn't necessarily help us predict the future. For example, here are two very different models that fit the current data pretty well:



The equation for “Model 3” is:

$$N = \frac{2000}{1 + e^{-0.32(t - 21.1)}}$$

It reaches its halfway point around March 21 and peaks at 2000 people infected. Unfortunately, “Model 4” also fits the data just as well:



Its equation is:

$$N = \frac{20000}{1 + e^{-0.24(t - 32)}}$$

But it reaches its halfway point at on April 1st and peaks at 20,000 people.

Reality could be anywhere in between, or even higher – I could have easily created a curve that fits the current data just as well and peaks at 2 million people. The point is that we just don't know because it all depends on how we act now.