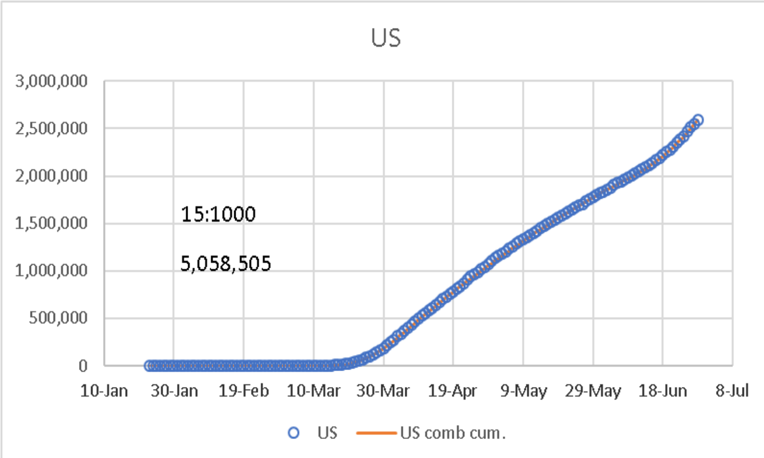
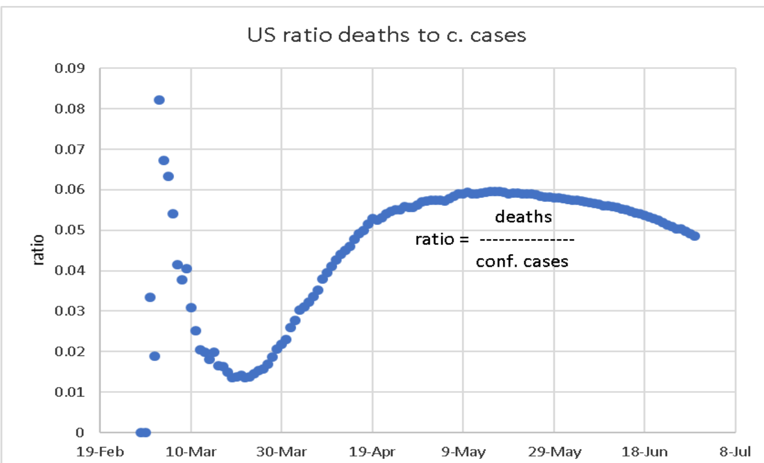
As of 30 June 2020, the US Covid-19 epidemic looks like this, cumulative:[1]



The ratio of deaths to confirmed cases looks like this:



So, it appears that a bit less than 5% of deaths are classified as due to CV-19, which shows this is no common flu, regardless of discussions on how these are classified.[2]

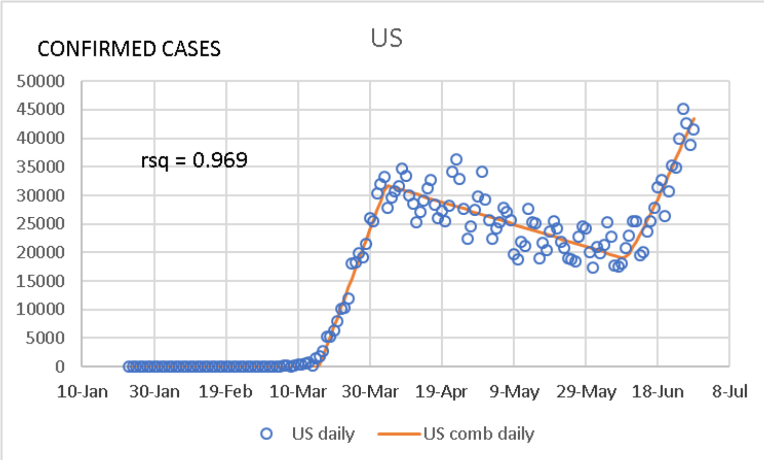
The distribution of this data could be applied locally, assuming the US case applies to the local situation in the Big Bend of Texas. This data has now been fitted with a combination of a triangular and a logistic, since it has taken a dramatic turn for the worse, and this offered a way to illustrate it. The Big Bend apparently lags in time the urban populations by quite a bit (see local graphs on this site), but we're obviously trying to catch up (no masks here and we like to party). But since an epidemic is nothing but a deadly diffusion phenomenon, sooner or later it reaches some saturation level that is hard to predict, often claimed to be 60 to 70% of the population.

---

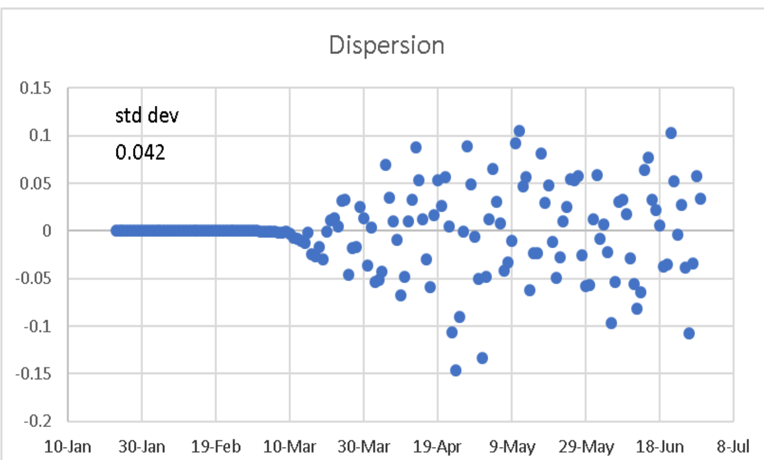
[1] https://github.com/CSSEGISandData/COVID-19

[2] The drop in the ratio could be attributed to more testing revealing more confirmed cases, though. Or better outcomes as treatments improve. Or both. As the cases ramp up, one would eventually expect to see this ratio begin to rise again, or at least slow its descent, unless this strain is a new one not as deadly. Not likely.

US appears like this, on a *daily* basis, not cumulative:

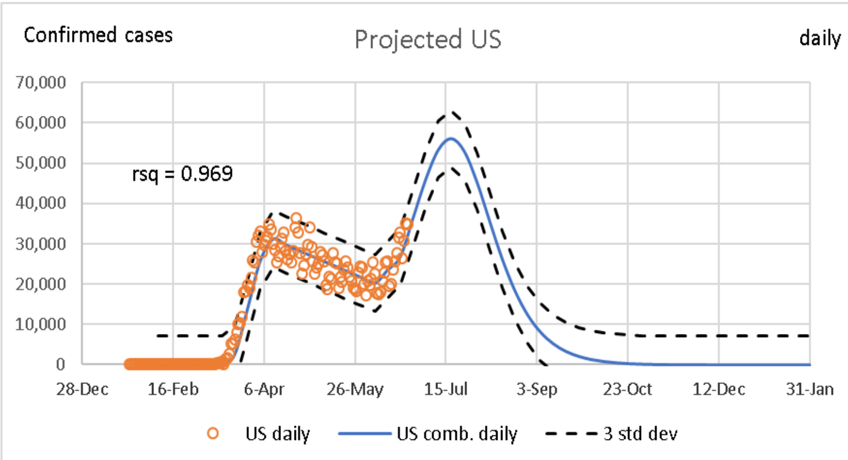


The curve fit is a hybrid of a triangular and logistic distribution. A measure of dispersion is the standard deviation of the residuals scaled to a relevant constant, such as the maximum value of the curve, or (y – y_hat) / y_hat_{max}. y is data recorded (in this case, the difference between each pair of successive values on the US version of the previously shown *cumulative* data), y_hat (pronounced *wye hat*) is the predicted value at each date of the curve fit, and y_hat_{max} is the maximum value of the distributed curve fit. The dispersion looks like this (below) and gives an idea of the variability of the data around the distribution. R^2 is for goodness of fit, shown as "rsq" on the graph, 1.0 being a perfect fit. Notice that you can have a nearly perfect fit still with a lot of dispersion. Or vice versa.[3]



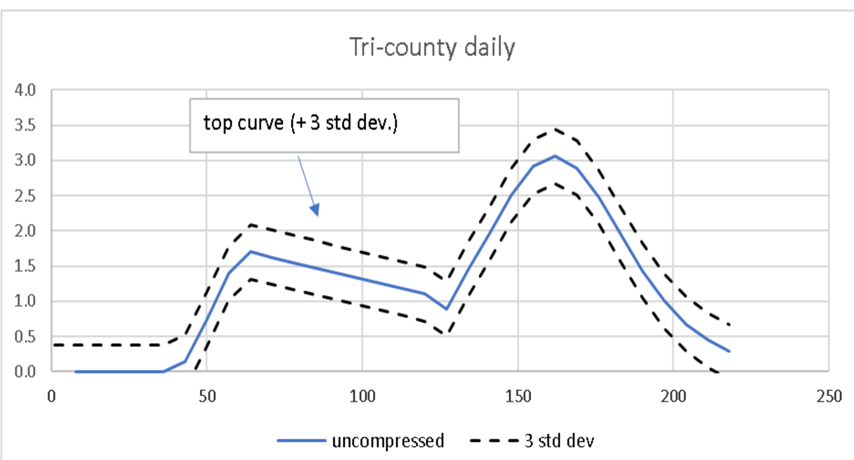
[3] As mentioned before, the gamma distribution was not fitting very well. That is why the data was switched to this hybrid.

The standard deviation of the dispersion values can be used to construct a 1 standard deviation boundary around the curve fit:

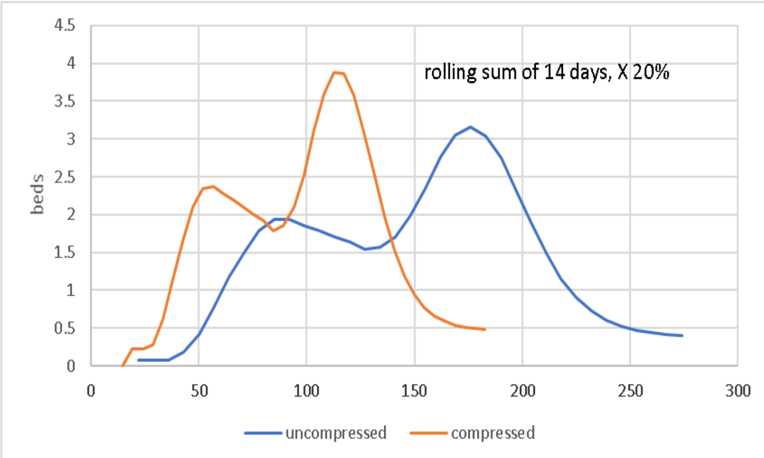


This has a double hump, and projects way off into the future. In any case, this data is for a population of 330 million then scaled to the Tri-County population of 18,000, in order to have an idea of what to expect here. The earlier comment on small sample size seems to be acting up in the Tri-County area—we have already jumped to 6:1000 range of confirmed cases, and still rising.

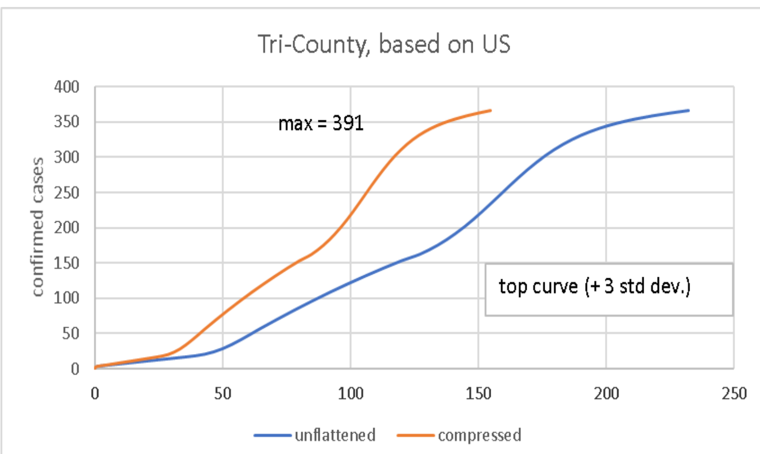
Below is the above US curve scaled to the Tri-County area, showing the top dashed curve as +3 std deviations above the base curve. What follows are some illustrations of what it looks like when such a curve is flattened (time scale is uncompressed, in other words). It is becoming a moot point on flattened curves—it's apparent any successful strategy has to go way beyond just flattening a curve.



This example is compressed by a factor 1.5; it's impossible to verify the real extent of compression, so this seemed like a reasonable assumption. Here is what that looks like:



The area under each curve is the same, there is no change in total cases. It's also worth looking at the cumulative version of the same data, to see what the time compression does to it. Notice it doesn't change the final, cumulative value.[4]



What seems apparent is trying to flatten a curve is an exercise in futility—what would make more sense is formulating a nation-wide strategy to tackle this (or at least state-wide), assuming no vaccine will ever appear. Lockdowns destroy jobs and the economy, so ought to be eliminated from the strategy.

When we were children, our parents taught us to cover our mouths when we coughed. It made sense then, and makes sense now—we didn't need Dr Fauci or the President to tell us to do it. At this point, we have an extremely active and dangerous virus, so it would make sense to step up the covering of the mouth by wearing a mask. You don't need a scientific study to tell you it works (although there are many), or refer to history (Asians have been doing this with viruses since the beginning of the twentieth

[4] If one assumes the Tri-County is a sample of the US population, scaling it this way will likely increase its variability, since that is the nature of small samples. The rolling sum probably smooths that out enough that it is not significant, but that is not addressed here. Tri-County infections finally kicked in, and appear very differently than this.

century) to just see the common sense of it: the velocity of the aerosol that comes out of your mouth is broken, just like when you cover your mouth, except more consistently. That way, the aerosol cloud that everyone creates by breathing, speaking, and coughing is much reduced in size. Furthermore, particle sizes are on the order of sub-micron size (viruses, and the larger water droplets they ride on), so the typical mask does *not* stop these at all from entering. So, no sense of false security, as many like to claim—assume there is *no* security against something passing through the mask to the wearer. It's all one way, from the wearer, to the rest of the world.

As far as Tri-County goes, the following tabulation shows what happens if the proposed US or Texas projections are scaled to our population:

max US daily	57,316	persons
equiv. max Tri-County	3.1	persons
max TX daily	9,180	persons
scaled to Tri-County	5.9	persons
max US cumulative	5,139,698	persons
equiv. max Tri-County	280	persons
incidence	16:1000	
max TX cumulative	751,773	persons
scaled to Tri-County	483	persons
incidence	27:1000	

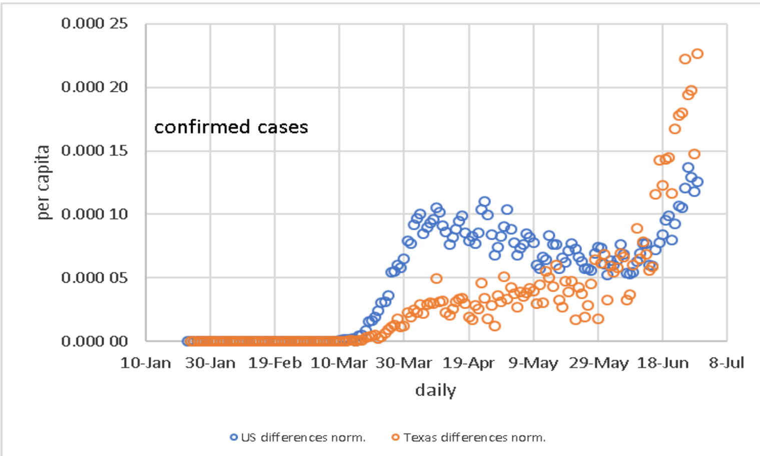
Tri-County has already busted through the highest daily number estimated above at 5.9 when it recorded 17 cases in one day. The cumulative amount is in the 6:1000 confirmed cases range, and still rising. You can look at the accompanying graphs to see what the rest of the world is experiencing. Some places have already got rid of it—why not here?

Shutdowns, masks, and isolation are really attempts to lower R_o,[5] but how much CV-19 is affected by lockdowns is not known well enough, and helps explain why the original introduction of lockdowns a couple of months ago only claimed to change duration and reduce maximum daily peak, not the total. An interesting experiment, probably influenced by the relative success of S. Korea, Taiwan, and New Zealand, is a state-wide regulation in Washington State, that anyone in public wears a mask, period. This obviously eliminates eating and drinking in public places, but this type of action may be what most contributed to the success of those countries listed. If there is enough compliance, it should be visible on the Washington graph pretty soon if it is effective. This may seem draconian, but would seem to have a much better chance of success than local county or city ordinances in actually lowering R_o, and not just simply flattening the curve while irritating everyone.

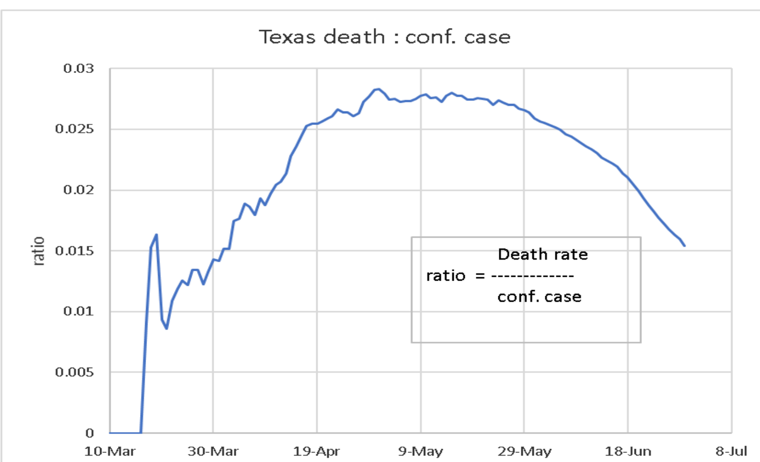
[5] Basic Reproduction Rate. R_o is a measure of transmissibility: $R_o < 1$, disease disappears; $R_o = 1$, it's endemic; $R_o > 1$, epidemic.

Postscript

If Texas and the USA data are plotted as daily data concurrently, on a per capita basis, here is what you see:

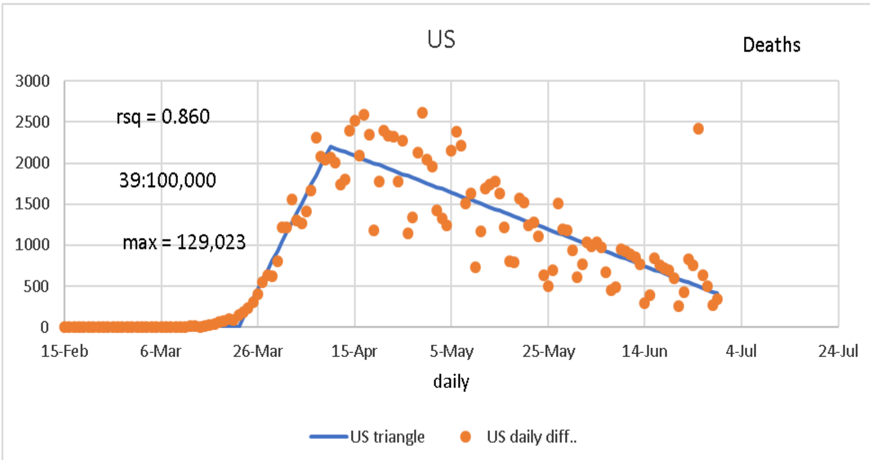


Texas is clearly going off the charts at this point. It's interesting that the Texas ratio of deaths to confirmed cases is dropping, and always has been significantly less than the US numbers, which could mean better treatment of a growing number of confirmed cases, or more testing finding more cases, or some combination of that. The good news, in a world of bad news, is a drop in the ratio might decrease the fear and panic factor and lead to more rational decisions by people in authority. Since deaths lag cases, we'll soon see what effect this latest strong spike has on deaths. Arizona has a similar curve.

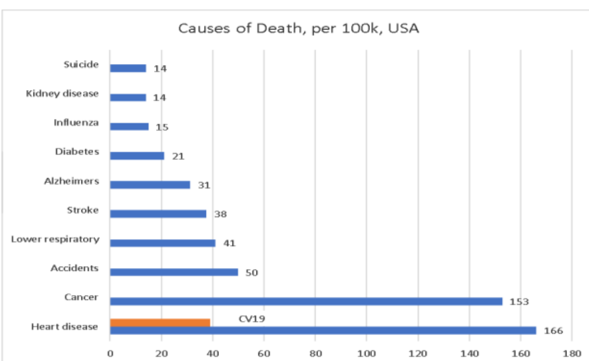


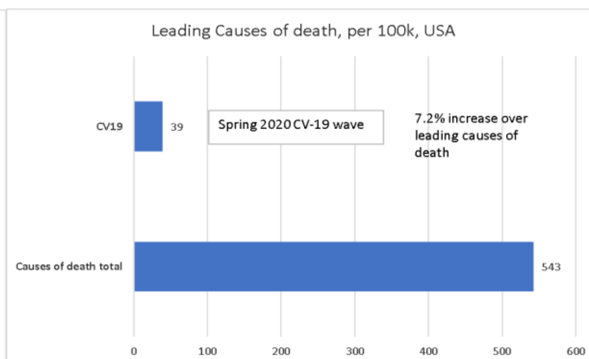
P.S.S.

A few weeks ago the Wall Street Journal published data on annual causes of death which I have plotted with the USA CV-19 death statistics. The USA CV-19 death statistic is based on a triangular distribution curve fit of USA deaths that generates a cumulative maximum for a completed wave.



This distribution predicts 129,000 deaths, once this particular wave is over. We may get one or two more waves this year. This value is plotted as deaths per 100,000 with the WSJ data:

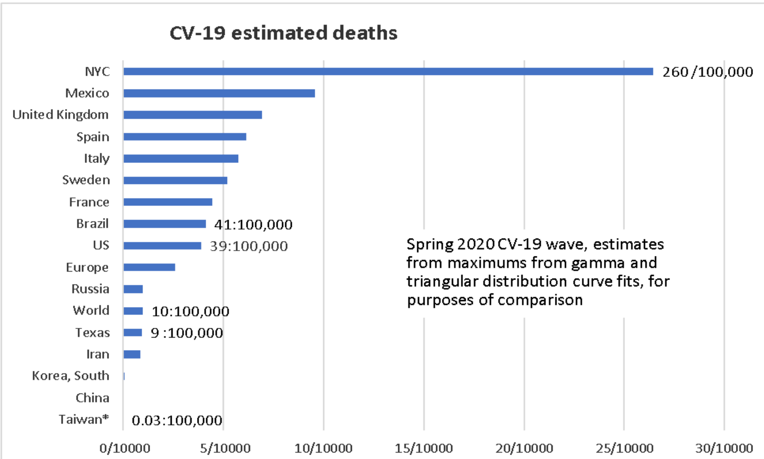




It's right up there with the major killers, although it's known many who died from CV-19 already had these potentially fatal issues. It's not clear where dying from old age fits in here, though. I suspect old age deaths would deduct from all these more or less equally, so the relative size remains unchanged.

NYC has suffered 260 deaths per 100,000 in this wave, while Texas still projects 9 per 100,000 (It's increasing, though.). Those are estimates based on the maximums predicted by their respective curve fits. It's something to keep in mind when considering the desperation found in NYC and the northeast United States; they have every reason to be desperate. On the other hand, Texas' incidence is still less than any of the categories listed above. It's still a serious issue for Texas, no doubt, but two orders of magnitude less than what NYC is experiencing, and one order of magnitude less than the US average. Certainly don't need to treat Texas like NYC, or, maybe we will have to, if people continue not to wear masks in public and to party like there is no tomorrow. Why do people need the Government to tell them to do what common sense should have told them already?

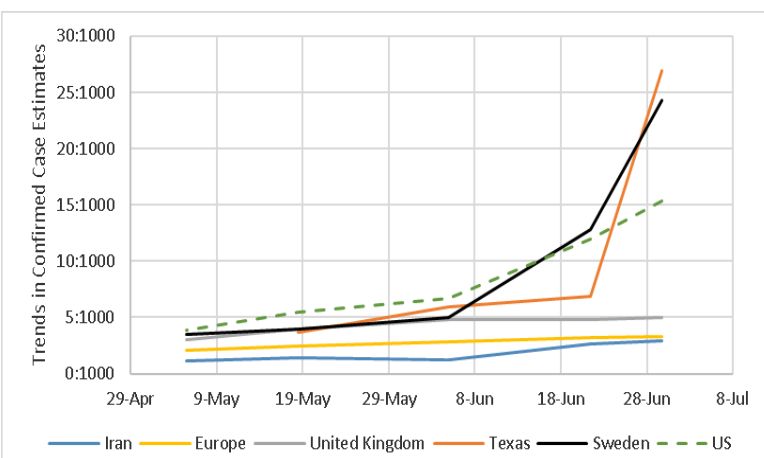
Using the estimator (the maximum value parameter from the gamma distribution curve fit, or the maximum cumulative value for the triangular fit), it's possible to do a bit of comparison of deaths around the world. Interestingly, two countries that did not do lockdowns (Sweden and Mexico) are mixed up in here. Mexico is not doing well at all, so far. I put NYC in here to show how much it skews the US data, since it is almost a fifth of all the cases. Not sufficient data to prove anything, but it does make one wonder, or I hope it makes one wonder.



For example, lockdowns are putting incredible stresses on people, losing jobs, being cooped up, maybe drinking too much. All of that stresses the immune system, and that makes people more susceptible to infection. In that case, lockdowns would not only *not* decrease the eventual number of infections, but might even increase them, due to the compromised immune systems.

Trends in Confirmed Case Estimates

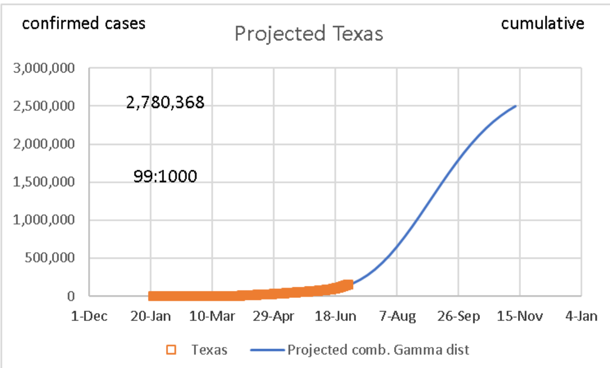
The confirmed case estimates are based on the maximum value parameter of a curve fit for a *completed* epidemic wave. The US was changed to a triangular distribution is being change to a combination curve, which predicts 20:1000 (not shown yet). Texas has such a severe spike it doesn't return believable results. The curve fit has been based on data that was in the process of being completed for the Spring 2020 CV-19 wave, so it's to be expected that the estimates might change. Here are confirmed case trends for a few select countries:

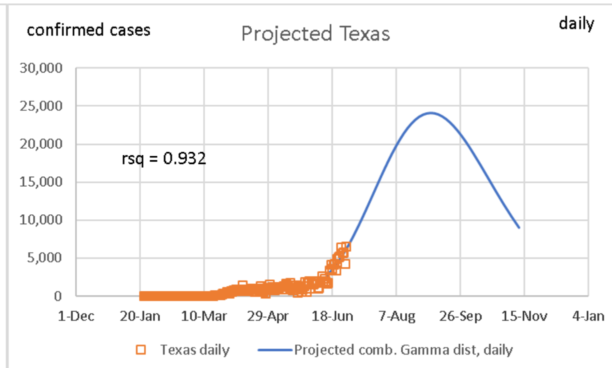


Interesting Projections for Texas

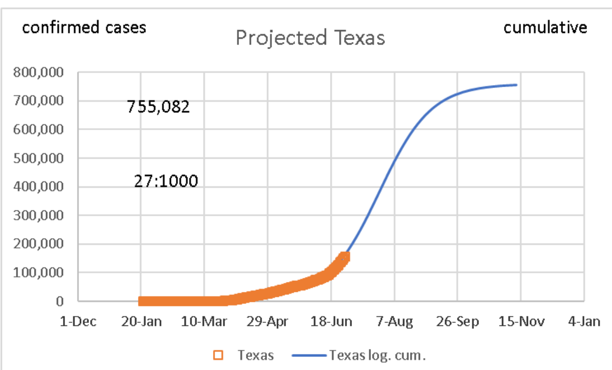
It's anyone's guess where the Texas numbers will end. Hopefully, corrective measures will be made that will include mandatory masks but no lockdowns. Contact tracing is becoming pretty prevalent, too.

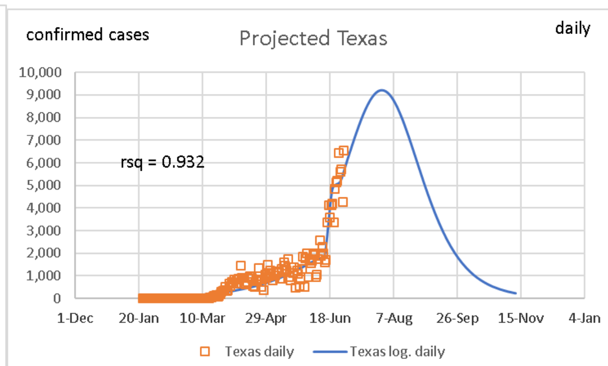
Gamma distribution projection:



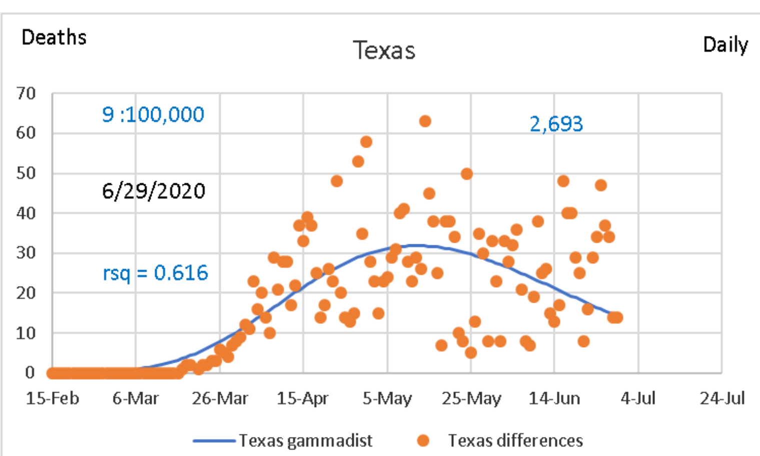


The logistic function predicts a much smaller number, but still large.





The sudden uptick that shows no sign of abating is certainly an issue for concern. So far, no corresponding increase in deaths—we hope it stays that way. Since the cases are primarily 20 to 30 year olds, perhaps that is the way it will go.

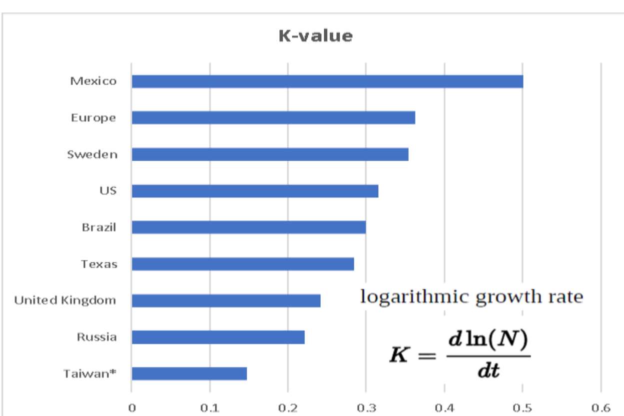


Explanation of statistic used for comparative purposes:

Often the generation of confirmed cases or deaths tend to be in a gamma distribution, which is a slightly skewed bell-shaped curve. Alternatively, the logistic distribution is used, as its parameters are less abstract then the gamma. Other times a triangular distribution better describes what's happening. A least squares procedure is used to get the closest fit to the data. Generally, the data is provided in a *cumulative* format, so the number increases each day, until the wave is over, at which point it is at its maximum and no longer changes. This can be converted to a daily format just by finding the difference between each pair of successive days. In mathematical terms, the daily format is the time derivative of the cumulative format. In any case, the area under the daily curve is identical to the last, largest and unchanging value in the cumulative format. This is the value that is used here as a comparative statistic, on a per capita basis. Each data set has a different beginning, develops at a different rate, so comparison of values on a particular date among any given datasets doesn't tell you much.

Cases are usually described as incidents per thousandths; deaths are described as incidents per hundred thousandths. In this way, one can get a better feel for the relative performance of countries, continents, cities, states, counties, as long as the current population figure is available. It's also a way to gauge one's own personal risk—for example, the projected rate for Texas of 27:1000 confirmed cases over several months is still a moderate risk for anything that's not fatal, despite being worse than average.[6] Keep in mind this statistic can and will change over time, not as a direct result of increasing cases or deaths, but because of the possible change in shape of the distribution. In some cases, particularly with the US, the distribution was changed from a gamma to a triangular because it fit better, which also changed the comparative statistic.

R_o is a pretty subjective parameter, but much discussed. If one can find the exponential part of the cumulative growth curve (a short section near the beginning of the cumulative curve, but starts to lose validity if less than N = 100 people), one can find the growth rate constant for an exponential, which is $K = d \ln(N) / dt$. (If plotted on semi-log paper, look for a straight line section near the beginning; K is the slope of that section.) $R_o = e^{K\tau}$ where τ is the average infectious period. Reduce K by reducing contacts, reduce τ by isolation of infected individuals, for example. Here are some K values:



[6] As far as fatalities goes, 9:100,000 mortality estimated for Texas is less than the 11:100,000 national rate for car fatalities. You have to ask yourself if you stay up at night worrying about getting killed in a car crash. Of course, those are all unconditional probabilities—given you're over 65 with significant health problems, it's a whole different story. And, eventually this estimate will likely increase for Texas.