

# The Effect of Temperature on the Spread of COVID-19

Shubh Goyal<sup>1</sup> and Melanie Malinas<sup>1</sup>

<sup>1</sup>Archbishop Mitty High School, San Jose, CA, USA

## ABSTRACT

This study analyzes the findings about the spread of COVID-19 in regions of varying temperatures. This paper will use the analysis of public data to conclude how impactful temperature is in determining the spread of COVID-19. In addition, this paper will shed light on some trends and some historical relevance to other pandemics similar in genetic makeup and growth. There have been numerous claims that higher temperatures will cause COVID-19 to diminish quickly and die out during the summer months. I wanted to test for seasonality as I would attain a better understanding of temperature's role. Other viruses like the flu and influenza are more common, and they display signs of seasonality. For example, the flu occurs in the Fall and Winter months which is why people get vaccines or flu shots in the preceding months. The conclusion about temperature's significance could also highlight the stage of severity of the virus. If the cases, start to drop for example, one could conclude that the virus is past its peak and will die out soon. Several states from the United States were monitored for their daily cases and temperature before being averaged into three-day intervals to depict a smoother trend. Both variables are public and accessible to all. The states were divided into colder states (New York, Pennsylvania, and Illinois) and hotter states (Texas, California, and Georgia). The results varied and displayed different trends. The colder states seemed to have gone past their peak and displayed a downward parabolic form. The temperature and cases displayed an inversely proportional relationship. The hotter states showed a constantly increasing case count with the graph showing a spike around the same time. The hotter states displayed an exponential trend and it is hard to predict when and how much they will peak. These two differing trends suggest that the effect temperature can have is not obvious across numerous areas. Therefore, temperature does not appear as a significantly impactful determinant in the growth of COVID-19.

The COVID-19 pandemic has rapidly spread from the city of Wuhan, China to almost every nation. COVID-19 is mainly spread by person to person transmission—mostly through respiratory droplets. It can stay airborne for over three hours, which means that the virus can be found in closed environments and can easily transmit from one individual to another (1). Another form of spread is through surfaces. By touching surfaces that have been in contact with individuals who have COVID-19, others can obtain the disease as well (2).

As of July 21, 2020, there have been approximately four million cases in the United States, and the number is constantly rising (3). To impede the spread of this virus, governments have asked people to stay a minimum of six feet apart, wear masks, and limit time spent outside. This should reduce the chances of obtaining the virus if someone were to sneeze or cough in a densely populated area. COVID-19 has originated from a strain of a coronavirus. There have been several coronavirus outbreaks in the past such as the SARS outbreak of 2002/2003 and the MERS outbreak of 2012. Corona translates to “crown-like” which is symbolic of the coronavirus strain's shape. These virions have crown-shaped spikes all over which allow them to bond to their hosts. COVID-19 bonds more easily to new hosts than SARS, even though it shares approximately 86% of its genetic makeup with SARS.

SARS and COVID-19 both affected a tremendous number of people, but they had different impacts. SARS displayed a nearly ten percent mortality rate in comparison to COVID-19's approximate three percent mortality rate. However, SARS did not spread as easily during its span of 2002-2003, so it ended up affecting about 8,000 people. COVID-19's mortality rate is lower; however, its ability to remain in the atmosphere and spread easily has accounted for its higher case count. It is hard to predict how many total people COVID-19 will affect or when it will end as scientists are still unsure about how to contain the virus. Possibilities such as a second wave, an evolved form, or seasonality could impact the growth of COVID-19.

One key aspect of SARS was its dependency on seasonality. SARS, influenza, and the flu are all seasonal, which means that they grow and develop during a certain time of the year (4). With COVID-19 being less than a year old, it is not possible to accurately predict how it will grow in the fall and winter. This is because seasonality is impacted by factors such as temperature and weather, but also human behavior. Humans are more prone to be outside during the summer months than the winter months. However, with these past viruses, scientists were able to develop trends between the temperature and the spread of the virus, allowing them to conclude that the flu develops in colder temperatures, hence winter (6).

## Experimental Procedures

The first step was choosing where to get the data. With varying cities, countries, and regions at one's disposal, there were a lot of potential areas to choose from. A unique aspect of the reaction to COVID-19 is that governments have instituted social distancing and lockdown, meaning that people can only interact in a socially-distanced manner or go outside for emergencies such as buying food. This paper utilizes data specific to the United States to ensure that the spread of the virus would not be affected by when lockdown was initiated. The United States has shown very high numbers in terms of the spread and infection rates. This would allow for a more accurate trend as the large amount of data would provide a more complete picture. Additionally, the United States is a large country with regions of varying temperatures, biomes, and populations.

The virus's growth can easily be impacted by the number of people in the environment. This brings the determinant of population into focus. A region with more people will allow for the disease to spread more easily and much more quickly than a region with fewer people. The population density also plays a critical role. If the population density is higher, then there will be more people per square foot, allowing the disease to spread easily from one individual to another. A decrease in population density would cause the virus to take longer to spread, which may cause it to die out before affecting as many people. To tackle the determinant of population and population density, this paper used data from the most populated and dense cities and states. The daily temperature was taken from the largest city in the state, and the daily cases were taken from the state. For example, the daily temperature of the New York dataset came from New York City; however, the daily cases were cumulative of the daily cases in all of the state of New York.

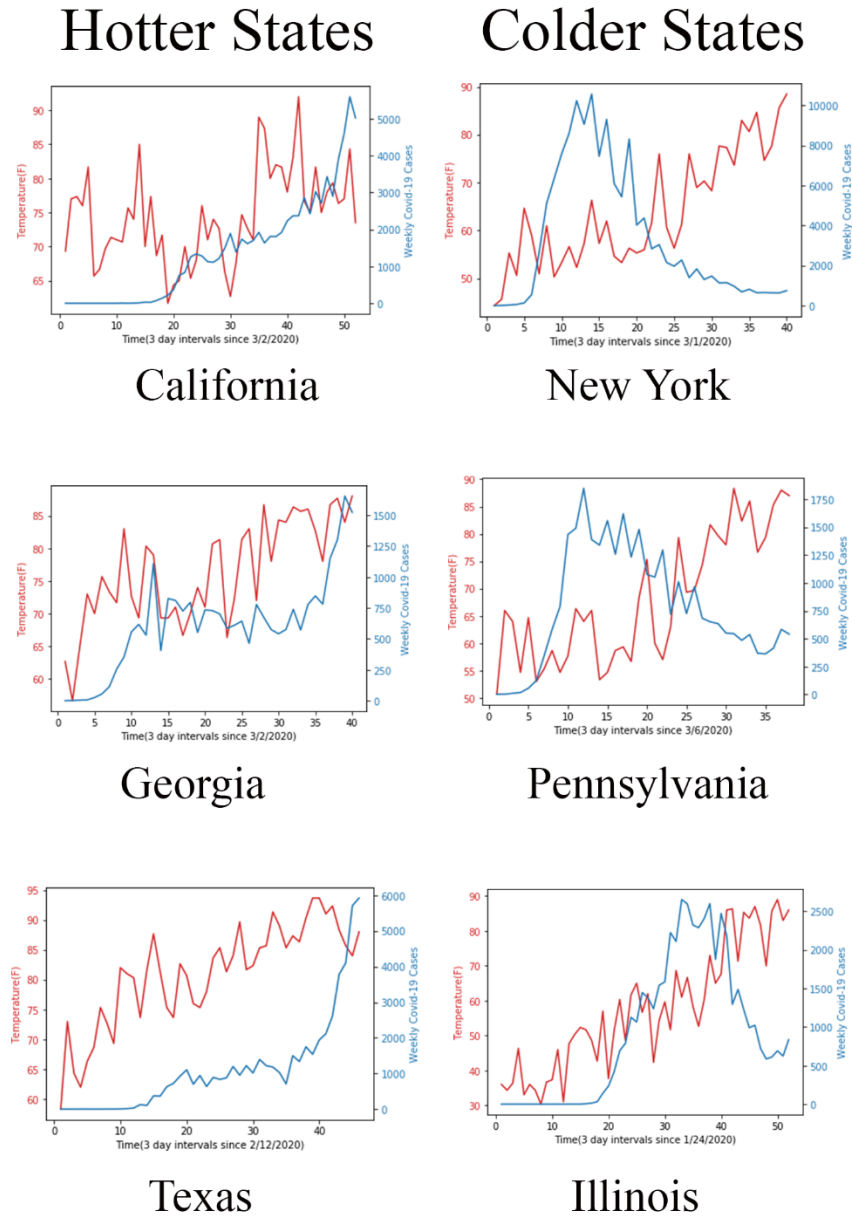
To obtain the data, the daily temperature was created in an array and taken from Accuweather.com. The cases were taken from the csv file in the public Github released by the New York Times. The Github data was public data released by the New York Times, so the data was ensured that it would be credible. The data was incorporated as a dataframe, and specific dataframes were created to display the cases of specific states.

1	date	state	fips	cases	deaths
2	2020-01-21	Washington	53	1	0
3	2020-01-22	Washington	53	1	0
4	2020-01-23	Washington	53	1	0
5	2020-01-24	Illinois	17	1	0
6	2020-01-24	Washington	53	1	0
7	2020-01-25	California	06	1	0
8	2020-01-25	Illinois	17	1	0
9	2020-01-25	Washington	53	1	0
10	2020-01-26	Arizona	04	1	0
11	2020-01-26	California	06	2	0
12	2020-01-26	Illinois	17	1	0

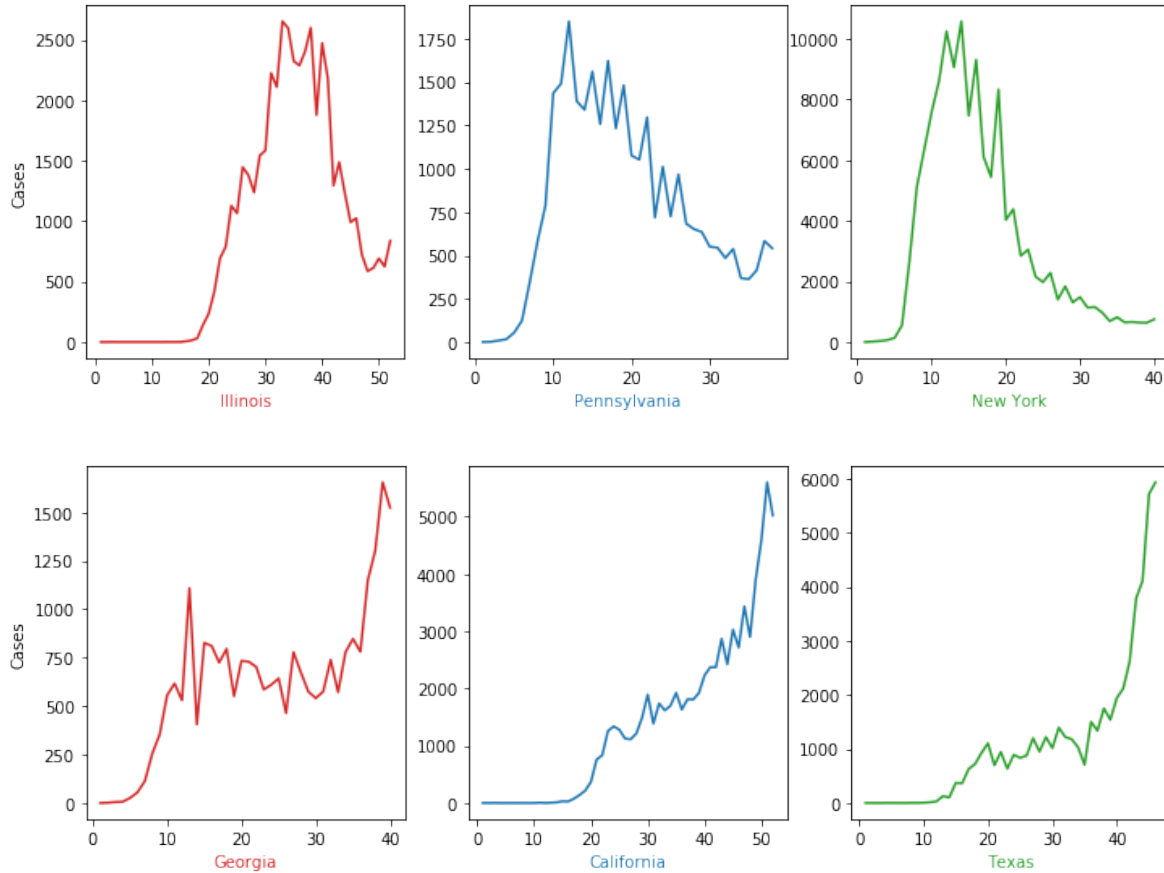
**Figure 1** shows an excerpt of the original data table of the Github. The figure displays the date, state, cases, and deaths. Instead of using the whole csv file, a locate command was used to filter the data to a key word such as the state name

in the state column. One can see that the cases are cumulative rather than representative of the daily cases. The code subtracted the cases of day D from the cases of day “D-1” to get the daily cases of day D, if D represented any random day picked from the dataframe. The fips column is used to locate the states on a map.

To create plots, a platform called Jupyter Notebook was used along with Python libraries such as Pandas and Matplotlib. The cases were in a dataframe which presented them as a cumulative case count monitored for each state at a daily interval. The x axis resembled days after the first case, and it was shared by both functions(daily cases and daily temperature). There were two y axes created, one for each function. This allowed the viewer to see both the number of cases as well as the daily temperature as time increased. In total, there were six states that were monitored, three hotter states and three colder states. The hotter states used were California, Texas, and Georgia. The colder states were New York, Illinois, and Pennsylvania. The x-axis was an array of days numbered consecutively until it was the same length as the state it was used for. The temperature and cases had a numbering system which was created by the software to match the range of values. Then, both functions were graphed on the same plot so that the reader could view both functions simultaneously. To graph figure 2, three subplots were created to depict the trend of each state. Then the colder and hotter states trend figures were merged to create one figure. See figure 2



**Figure 2** shows the plots of the six different states. The left column displays the hotter states, and the right column shows the colder states. The start of the x-axis is the first case for that state. Each state depicts two functions with the red depicting the temperature and the blue depicting the COVID-19 cases. The daily temperature and cases were averaged into three day intervals to reduce the number of spikes and plot a smoother trend.



**Figure 3** shows the differing trends of the colder and hotter states. The colder states above show a parabolic trend. They are past their relative maximum and appear to be decreasing. However, the end does show a small increase in cases, and this could be a small spike or possibly a start of a second wave. The hotter states below show a more exponential trend. The hotter states do show a bit of a decrease towards the end. This may be a sign that it is following the parabolic trend, or might just be a small decrease insignificant to the larger picture.

## Limitations

This study came with some limitations. One limitation was the timing of social distancing and lockdown implementations. Certain states closed and opened earlier which varied the severity of cases. This probably altered the results, because if the virus's peak occurred during rather than before lockdown, the number of people affected would be different. Also, the daily temperature was monitored from the most populous city in that state. This is because the temperature would vary in different parts of each state, which could slightly skew the results.

## Discussion

Each of the individual state graphs display their own story. The California graph shows the daily temperature spiking up and down constantly without much change, yet the number of daily cases constantly increases. Unlike the other graphs, the daily temperature and daily cases functions for Texas rarely overlap, but the daily cases still increase as

temperature increases. The temperature and cases functions of the Georgia graph display a similar path as they both also show a positive correlation. The New York graph shows a strong inverse relationship between the cases and temperature with the cases converging to zero. The Illinois graph initially supports a positive correlation, yet in the latter half the two functions diverge as temperature increases and the daily cases decrease. Like New York, Pennsylvania also shows a strong inverse relationship. The states, when grouped, show a similar trend. The hotter states all display an exponential type trend where the cases are gradually increasing and then suddenly spike up after about 120-140 days. The colder states show a more parabolic trend as the cases start to diminish as the temperature continues to increase. These trends can be shown in Figures 7 and 8. The original hypothesis was that the states would show an inversely proportional relationship between temperature and daily COVID-19 cases. The colder states displayed this effectively, but the hotter states did not. It is possible that the hotter states will eventually converge as the virus diminishes and dies out, displaying a parabolic shape like that of the cold states. Overall, temperature does seem in the span of a couple months to affect the growth of COVID-19, yet it does not display a strong overarching association. This may be due to the effect of other factors, or temperature may not be a strong enough determinant to massively impact the spread of COVID-19.

## References

Nazario, Brunilda. "How Coronavirus Is Transmitted: Here Are All the Ways It Can Spread." *WebMD*, WebMD, 20 July 2020, [www.webmd.com/lung/coronavirus-transmission-overview](http://www.webmd.com/lung/coronavirus-transmission-overview).

"Novel Coronavirus 2019, Wuhan, China." *Cdc.Gov*, 2020, [www.cdc.gov/coronavirus/2019-ncov/index.html](http://www.cdc.gov/coronavirus/2019-ncov/index.html).

Boice, Ryan Best, Jay. "Where The Latest COVID-19 Models Think We're Headed — And Why They Disagree." *FiveThirtyEight*, 1 May 2020, [projects.fivethirtyeight.com/covid-forecasts/](https://projects.fivethirtyeight.com/covid-forecasts/).

"Here Are the States Where COVID-19 Is Increasing." *Healthline*, 21 July 2020, [www.healthline.com/health-news/here-are-the-states-where-covid-19-is-increasing](http://www.healthline.com/health-news/here-are-the-states-where-covid-19-is-increasing). Accessed 21 July 2020.

"Nytimes/Covid-19-Data." *GitHub*, [github.com/nytimes/covid-19-data/blob/master/us-states.csv](https://github.com/nytimes/covid-19-data/blob/master/us-states.csv).

Ujiie, Mugen, et al. "Effect of Temperature on the Infectivity of COVID-19." *International Journal of Infectious Diseases*, vol. 95, June 2020, pp. 301–303, 10.1016/j.ijid.2020.04.068. Accessed 11 May 2020.

"Some States Are Learning What Happens to COVID-19 Cases If You Reopen." *Healthline*, [www.healthline.com/health-news/covid19-cases-rising-states-reopened](http://www.healthline.com/health-news/covid19-cases-rising-states-reopened).