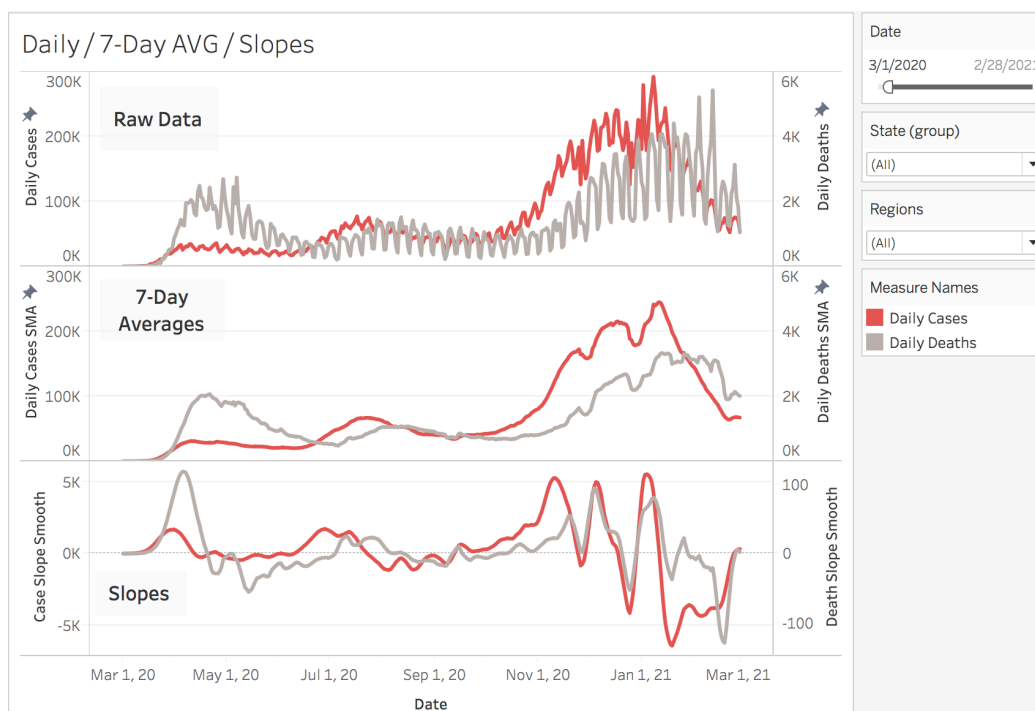


## Tableau visualizations of COVID-19 data

The COVID-19 pandemic has been at the forefront of the nation's attention since early 2020. For a data scientist, the intense scrutiny of the pandemic and resulting data is an oasis of opportunity; there are very few examples of large public datasets with such significant ramifications. Using compiled and curated state-level data in combination with Tableau, this writing sample will examine the utility of visualizations in understanding datasets and deriving actionable information from them.

Creating well-formatted line graphs can provide excellent insight into the general trends of cases and deaths across the nation. The following example shows three separate line graphs which reveal noteworthy aspects of the dataset.



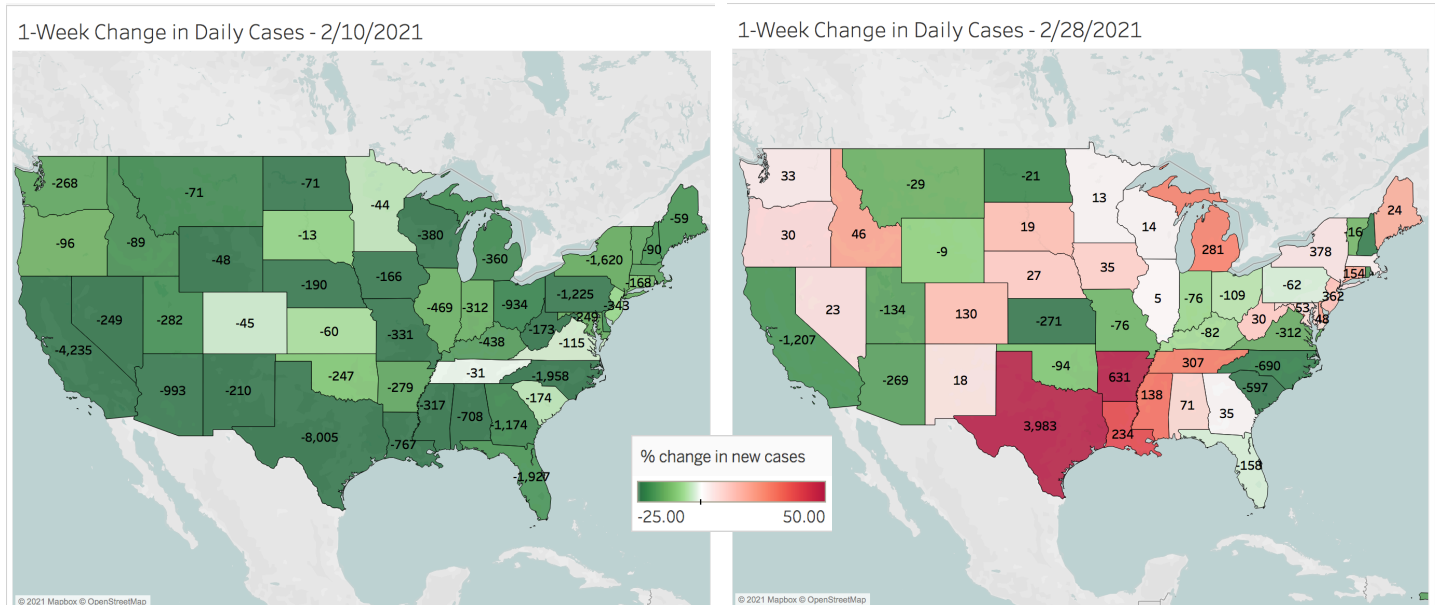
In the top graph, there is a very apparent periodicity. This periodicity occurs on a weekly basis, and is a result of data being reported unevenly on different days of the week. In order to smooth out the raw data and gain a better understanding of the trends in the data, the second graph shows the 7-day simple moving average (SMA). This graphic is generally preferable due to its increased legibility and less extreme variation. Using this second graph, three local maximums are clearly visible centered around April 2020, July 2020, and January 2021.

The final graph shows a smoothed derivative (slope) of the second graph. This graphic is useful in assessing how quickly things are getting worse or better. For example, the final points on this graph show both the cases (red) and deaths (grey) at or just above a slope of zero after a lengthy period of negative values. This indicates that the improving numbers the U.S. has been

## Tableau visualizations of COVID-19 data

enjoying since mid-January have come to an end. As a nation, we may be beginning a new period of case-load stability, necessitating further mitigation strategies or vaccinations to see continued improvement.

Another valuable tool for data analysis in Tableau is interactive graphics. Using Tableau's integrated mapping tools, we can create an interactive geospatial representation of case trends. The final element to this visualization is calculated fields which indicate the rate of change as a percentage, as well as a daily case value vs. 1 week prior. The simple way to explain this is: "Have things gotten better or worse in the last week?" In the first graphic (left) from early February, all of the lower 48 states are enjoying decreasing caseloads as the US recovers from the holiday spike. Moving ahead to the final day of February though, rising numbers of COVID-19 infections can be seen across the country, especially in the South.



Key contributors to the rising case load are clearly visible in this graphic, including Texas, which had an average of nearly 4000 more daily cases in the week ending on 2/28 than the week prior. Unlike the daily case values which are centered in each state, the shading of the state is on a percentage scale. This allows the graphic to indicate both absolute change and relative change within the same visualization. This is the reason some states with less infections like Maine and Idaho appear in a darker shade of red than New York, despite having much lower absolute case increases. Combining these measures into a single visualization allows the viewer to gain a strong, multifaceted understanding of the situation in a relatively short time.