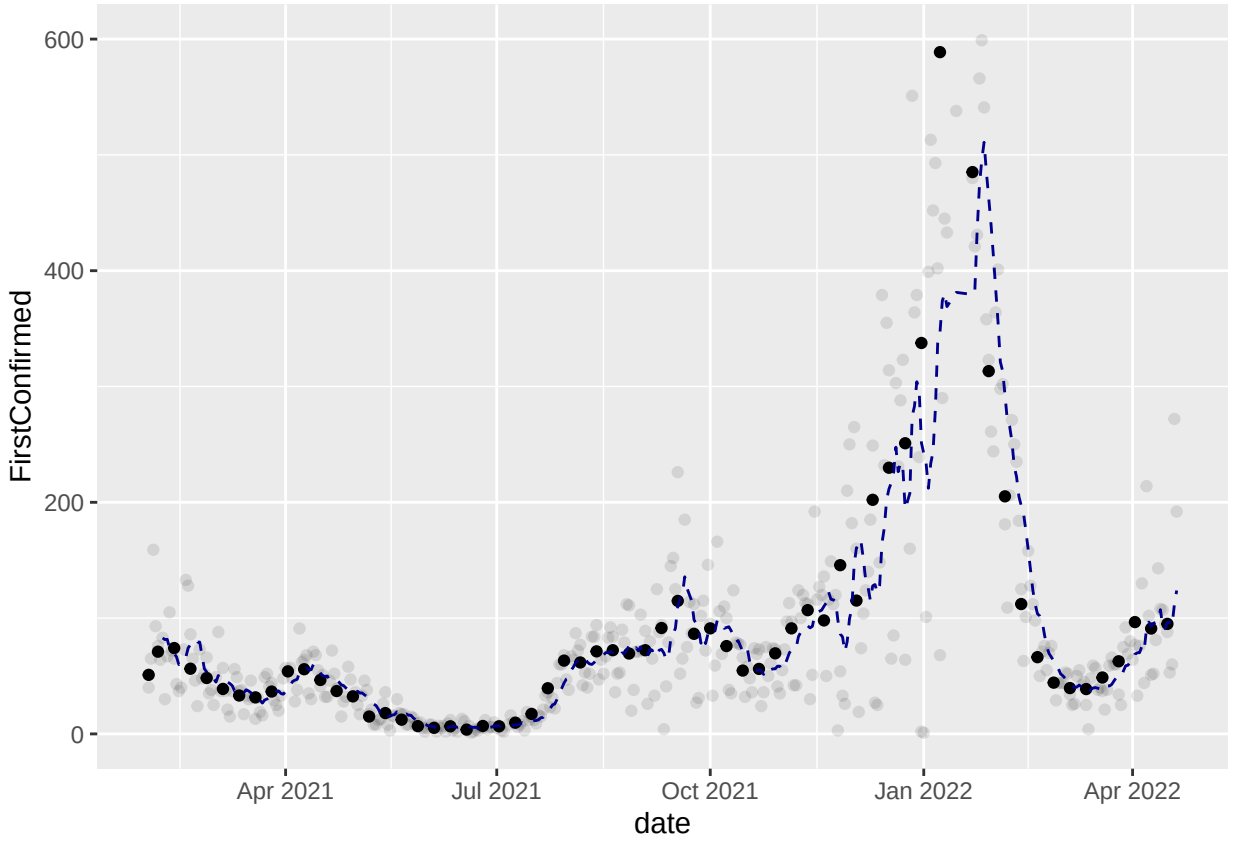
Does a Day-of-the-Week effect exist?

Marlin Lee, Steve Goldstein, Kyllan Wunder, Abe Megahed
University of Wisconsin Data Science Institute - August, 2022

7 Day smoothing

When a 7-day smoothing is applied to the case data it is apparent that the data does not always follow the line perfectly. When we add in the average of every week we can see that many values are not close to the weekly average at all.



Cases per day

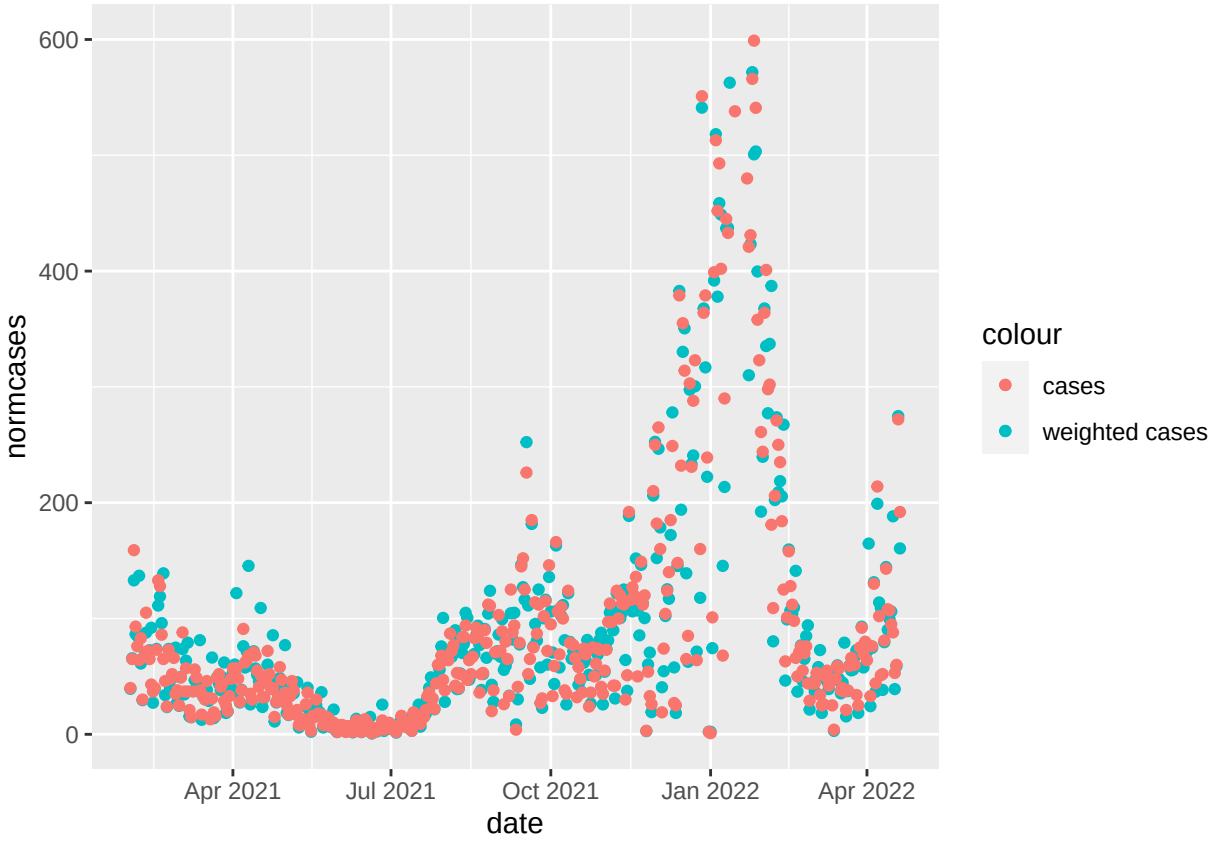
Looking closer at the data if we sum all of the cases by the day of week in which they were reported and find the proportion of how many cases land on each day we can clearly see that Sunday has the highest proportion of cases confirmed followed by Wednesday, while Friday has the lowest proportion. To try and combat this issue we have created a normalizing factor based off of the proportions for each day of the week. This is just the reciprocal of the proportion.

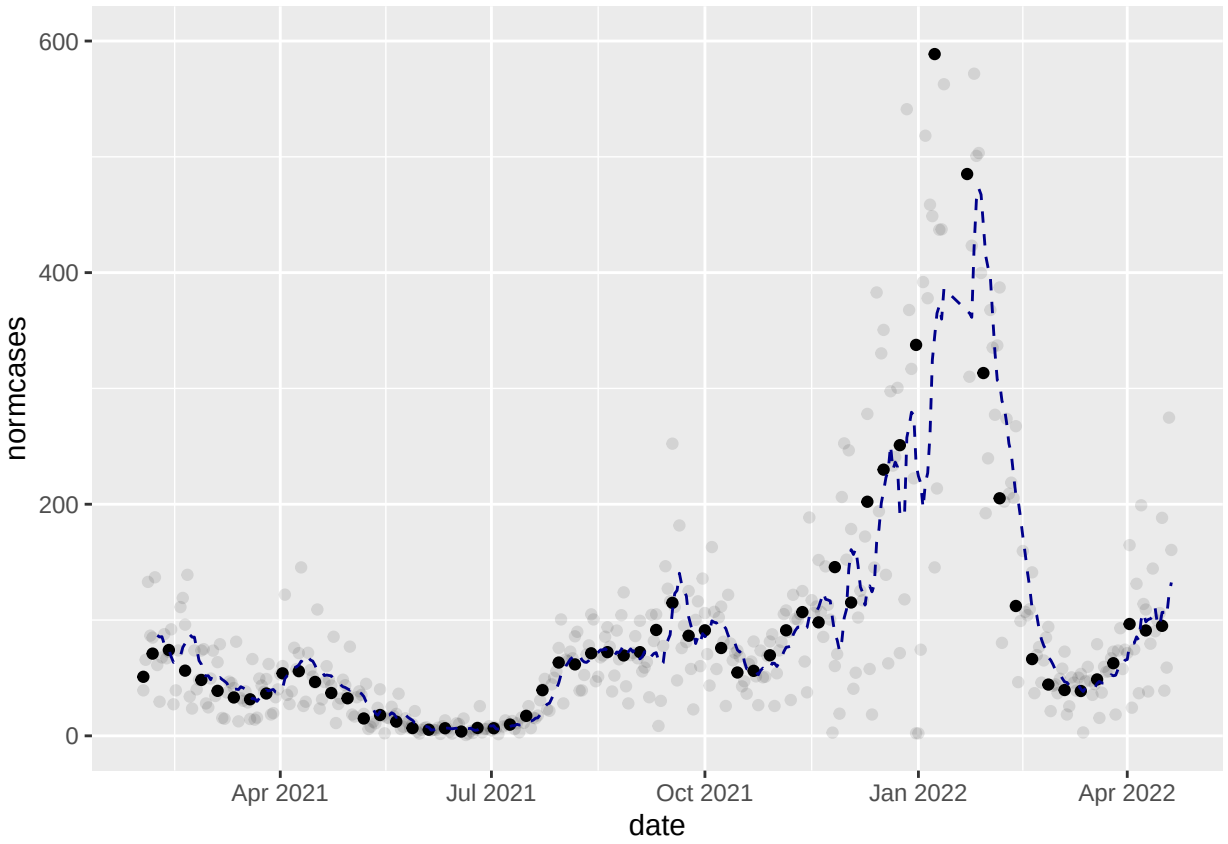
```
## # A tibble: 6 x 5
##   dayofweek sumperday     n  prop normalizingfactor
##   <ord>         <int> <int> <dbl>             <dbl>
```

```
## 1 Sun           10528     1 0.194           0.737
## 2 Mon            7896     1 0.145           0.982
## 3 Tue            7676     1 0.141           1.01
## 4 Wed            9275     1 0.171           0.836
## 5 Thu            8335     1 0.154           0.930
## 6 Fri            6946     1 0.128           1.12
```

Normilizing cases

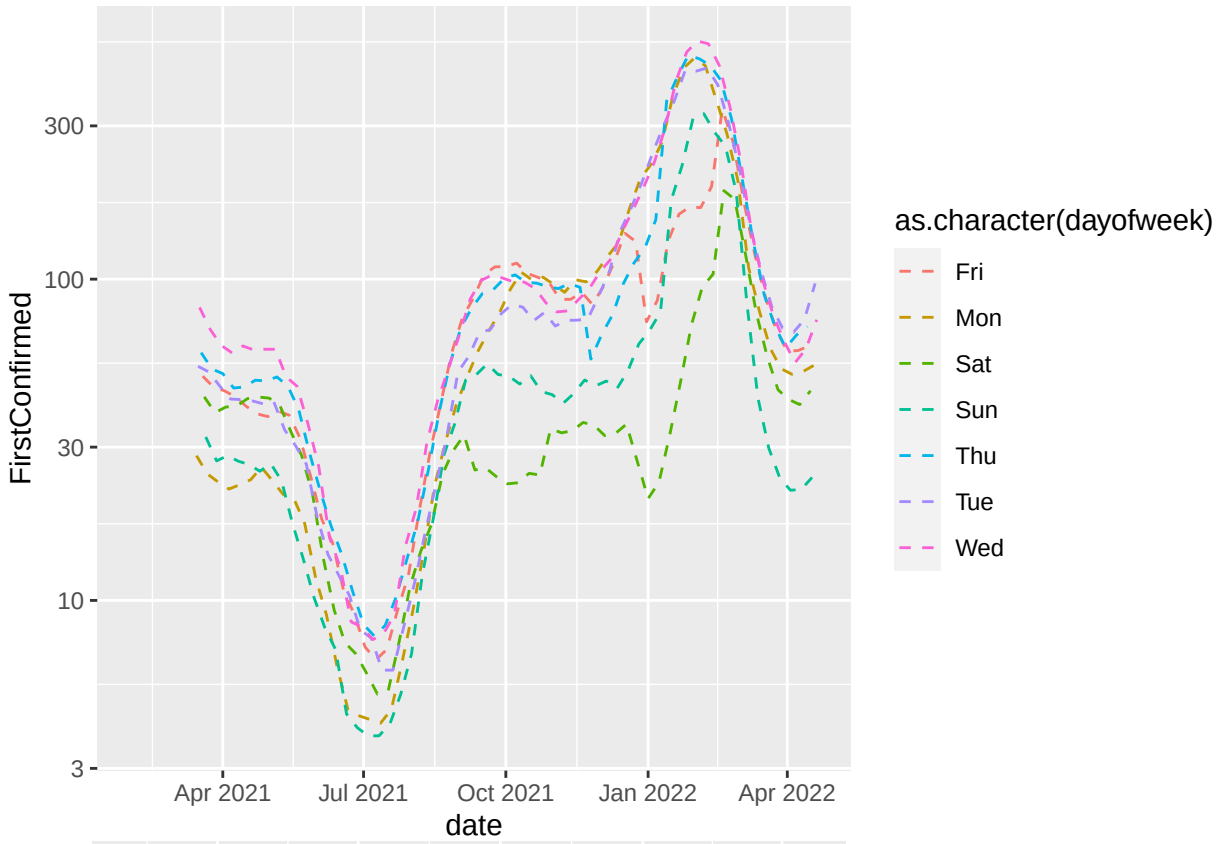
As shown normalizing cases is strange. In some places cases are shifted up and some are shifted down. It is hard to tell weather if this has helped or not. But as shown using the same graph from before with the weekly averages not being changed clearly the cases have shifted a lot.

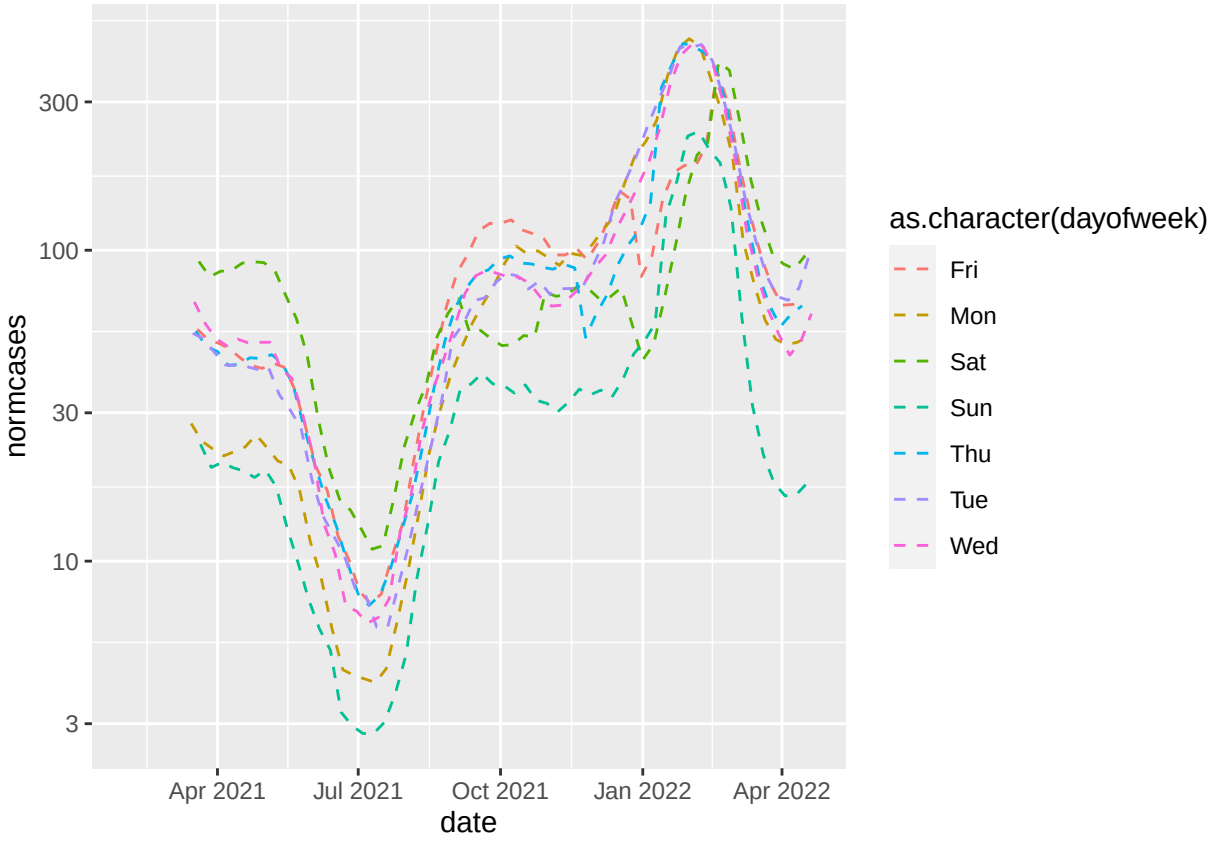




Does normlizing help day-of-week effect

If we plot out number of cases per day we can see that some days are always higher like Wednesday (Pink) and some are always lower like Saturday (Green) and Sunday (Light Green). After normalization while Saturday and Sunday are still usually the lowest the rest of the days converge and come closer together and less distinct. It is also clear to see that the days that have the highest and lowest proportions change over time leading us to belive that a better normalizing factor is needed as the proportion of cases changes.


300
100
30
10
3
FirstConfirmed
Apr 2021
Jul 2021
Oct 2021
Jan 2022
Apr 2022
date
as.character(dayofweek)
Fri
Mon
Sat
Sun
Thu
Tue
Wed


300
100
30
10
3
normcases
Apr 2021
Jul 2021
Oct 2021
Jan 2022
Apr 2022
date
as.character(dayofweek)
Fri
Mon
Sat
Sun
Thu
Tue
Wed