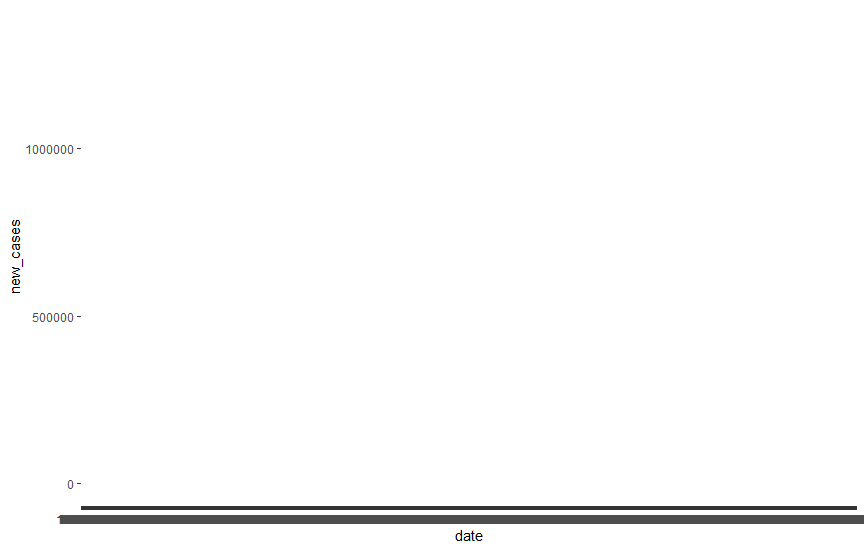
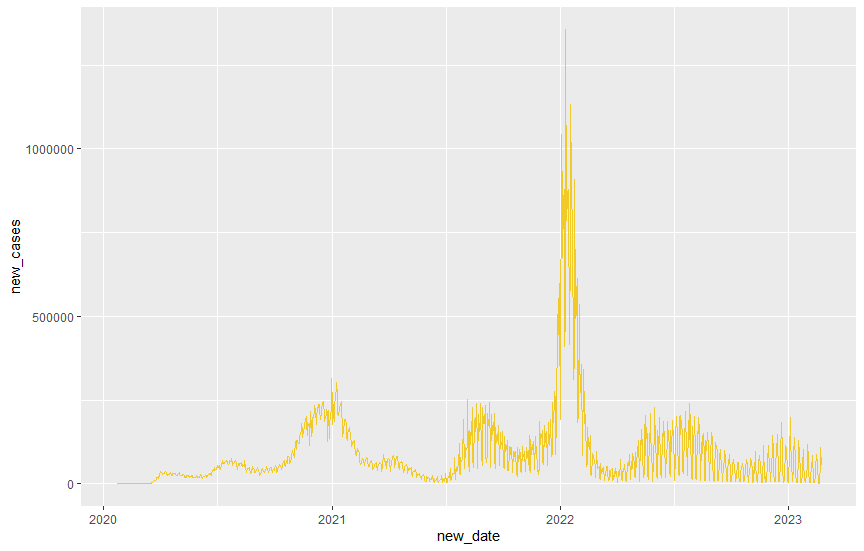
***HW 6: Use the provided data at the website (***[***https://ourworldindata.org/coronavirus/country/united-states?country=~USA***](https://ourworldindata.org/coronavirus/country/united-states?country=~USA)***)*** ***or All of Us data to complete the following visualization of amounts. Choose a state or region in the US and display the amounts of either hospitalizations for COVID-19 or Coronavirus infections. I am not going to assign a specific region or state or whether you choose hospitalizations or infections since I want everyone to create a unique dataviz. I will not grade you on your selection of region, but instead the quality of the dataviz.***

I am interested in daily new cases and daily administered vaccines which are fields “new\_cases” and “new\_vaccinations” respectively. First, I want to know how the new cases look over time and then I want to add the vaccines variable to see if the cases started declines or increasing.

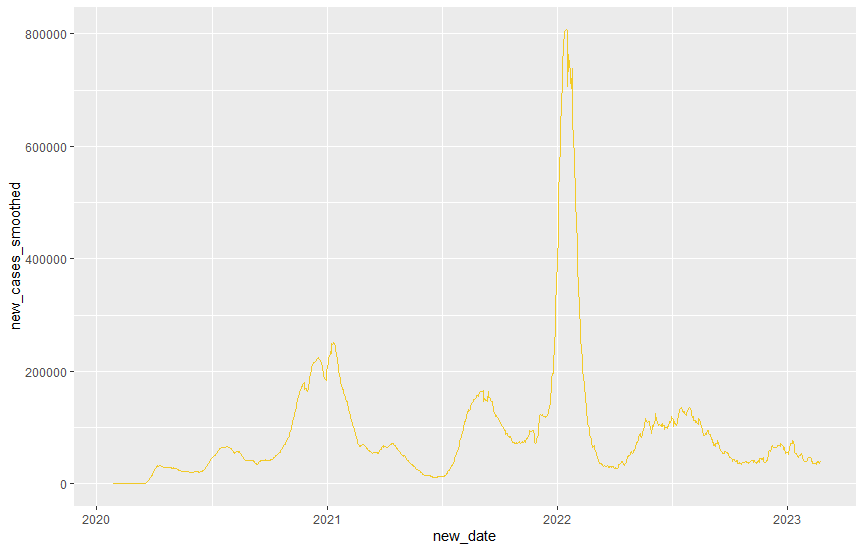
My first attempt was to do a very simple default plot of new cases and this was the result:



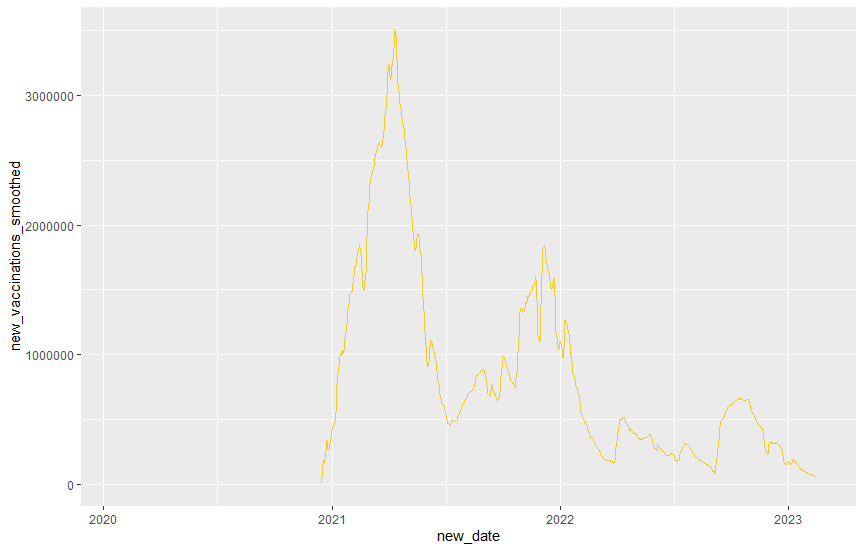
As we can see the 1000+ values of Dates in the x-axis make the graph very hard to read so the first thing I did was to scale the x-axis according to Year. Like this:



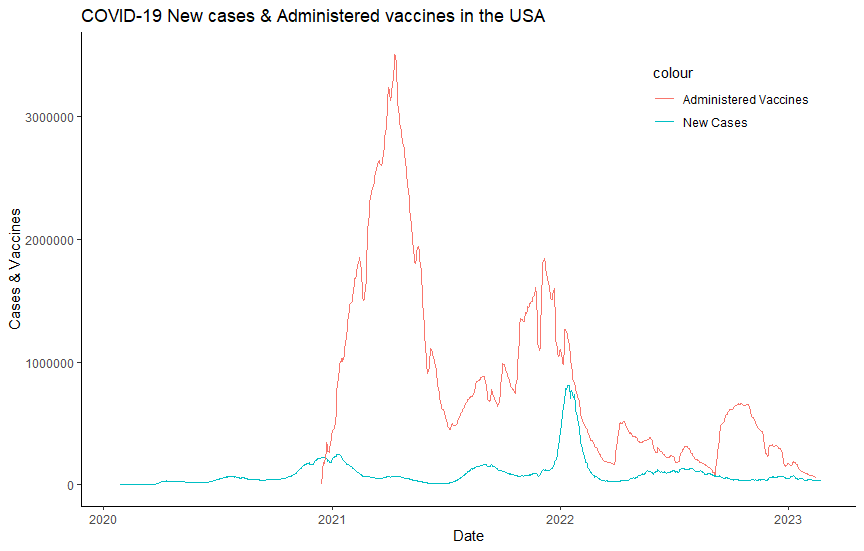
I now noticed the line was broken and realized that my original dataset contains the same variable but smoothed. Using it now the graph looks like this:



Next, I produced a plot for new vaccines using the smoothed version.



Finally, I plotted the two variables in the same plot using different colors to distinguish them. Additionally, I modifying the theme, labels, plot title and both axis titles.



With this visualization I can have more information and conclude that although vaccines has been administered for around a year from 2021 to 2022, the start of 2022 saw a considerable spike in new cases. Possible reasons: New virus strains with higher infection rate AND less movement restrictions to the population, hence, more chance for contagion.

Future work: I want to use the “stringency\_index” field however, it is a percentage field and I found it hard to incorporate it into my plot. Maybe use a 2-plot approach or add a second axis for this new field. I think it would be valid because it is a percentage so it might look good at the top of my current 2 variables. I did not do this because R’s ggplot2 doesn’t allow you to add a second y-axis with values from a variable.