# Report

Juehan Wang

12/7/2021

## Introduction

As of November 30th, the coronavirus disease 2019 (COVID-19) pandemic had resulted in 49,716,825 cases and 806,398 deaths in the United States. As of November 24, 2021, around 28 percent of total COVID-19 deaths in the United States have been among adults 85 years and older, despite this age group only accounting for 2 percent of the U.S. population.The widespread availability of vaccines in the United States helped to reduce cases in the spring and early summer. In this analysis, we obtain the COVID-19 Deaths and Fully Vaccinated Status Data, collected from January 1st, 2020 to December 1st, 2021. The main question that this analysis explores is the association between the number of deaths due to COVID-19 and the number of people who have been fully vaccinated in different age groups.

## Methods

Three raw datasets are included in this analysis, which are COVID-19 deaths data by state and age group, COVID-19 fully vaccinated data by state and age group and state population data. COVID-19 Deaths Data involves corona virus disease 2019 (COVID-19) and pneumonia reported to NCHS by jurisdiction of occurrence, place of death, and age group, collected from 01/01/2020 to 12/01/2021. The link is https://data.cdc.gov/NCHS/Provisional-COVID-19-Deaths-by-Place-of-Death-and-/4va6-ph5s. COVID-19 Vaccinations Data includes the overall US COVID-19 Vaccine administration and vaccine equity data at county level, collected from 01/01/2020 to 12/01/2021. The link is https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh. Population of 2019 Data involves the estimates of the Total Resident Population and Resident Population Age 18 Years and Older for the United States. https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-detail.html.

To analyze the data, R packages, dplyr, data.table, tidyr, tidyverse and usmap, were used.

In order to get the final analysis data, first, read in the data. Second, clean the raw datasets. For COVID-19 deaths data, the variables we are interested in counts of COVID-19 deaths by state and age group. Therefore, we clustered age groups into three groups, 0-18, 18-65 and 65+. Then the proportions of COVID-19 deaths among all deaths were calculated. Abbreviated state names were combined to the data. For COVID-19 fully vaccinated data, the variables of interest are the counts of fully vaccinated people by age group and state. Full state names were combined to the data. For state population data, the variables that we are concerned about are the population of different states and age groups. Therefore, we clustered age into three groups, 0-18, 18-65 and 65+. Third, merge the datasets and do some cleaning. Then we check the dimensions of the datasets. The COVID-19 deaths data and the fully vaccinated data have same number of columns. The population data have one more state data than the other two datasets. So, it should be mentioned here that the extra data will be deleted during data merging step. After checking the heads and tails and making sure that the datasets have good dimensions, we merge the datasets and do some cleaning, and then we get the final data for analysis.

As for statistical methods, descriptive analysis was done by summarizing statistics of the variables that this study concern about. Plots were shown by bar charts and maps, in order to have a straight forward view of the concerned variables. The main question of this study was explored using correlation analysis and smooth graph, based on proportion data.

# Results

## Preliminary Results

Bar charts and maps are shown to have a better view of the condition of COVID-19 Deaths and status of fully vaccinated in different states among age groups. From the bar chart of death due to Covid-19 we can straightly get that the largest numbers of deaths due to Covid-19 were in Texas and California. It is also obvious that the numbers of deaths in people older than 65 were much larger than those in other age groups. When looking at the bar chart of death proportions due to Covid-19 among all deaths, we get that the largest proportion was in Texas. From the bar chart of numbers of people who were fully vaccinated we can straightly get that the largest number of people who have got fully vaccinated was in California. It is also obvious that the numbers in people who were between 18 to 65 years old were much larger than those in other age groups. When looking at the bar chart of fully vaccinated proportions among the population, we get that the largest proportions were in Vermont, Rhode Island and Connecticut. Also, the proportions of fully vaccinated in people who were older than 65 years old were larger than those in other age groups. In the maps of deaths, states that are darker in the maps have more deaths or larger proportion of deaths. In the maps of status of fully vaccinated, states that are darker in the maps have larger number or proportion of people who were fully vaccinated.

## Final Results

The association between COVID-19 Deaths and fully vaccinated status is explored by correlation analysis of the proportion of COVID-19 deaths in all deaths and the proportion of fully vaccinated people in the population. The correlation between these two variables is 0.02, which is close to 0. Therefore, these two variables are not related. The association between COVID-19 Deaths and fully vaccinated status in different age groups is also showed in this smooth graph. The plot shows that there is no definite or fixed trend in the relationship between the two variables. There seems to have an increasing trend when the fully vaccinated proportion is small, but a turning point appears when the proportion increases. Still, the confidence intervals are quite wide as shown in the plot, so the trends are not reliable.

# Conclusion

There is no association between COVID-19 Deaths and fully vaccinated status in different age groups based on the current data. It is also found that people over the age of 65 have the highest deaths, which means that younger people are more likely to survive from COVID-19. Meanwhile, people older than 65 years old have the largest proportion of having been fully vaccinated among the whole population. In future analysis, longer term association between deaths due to COVID-19 and the status of fully vaccinated among different age groups should be looked at.
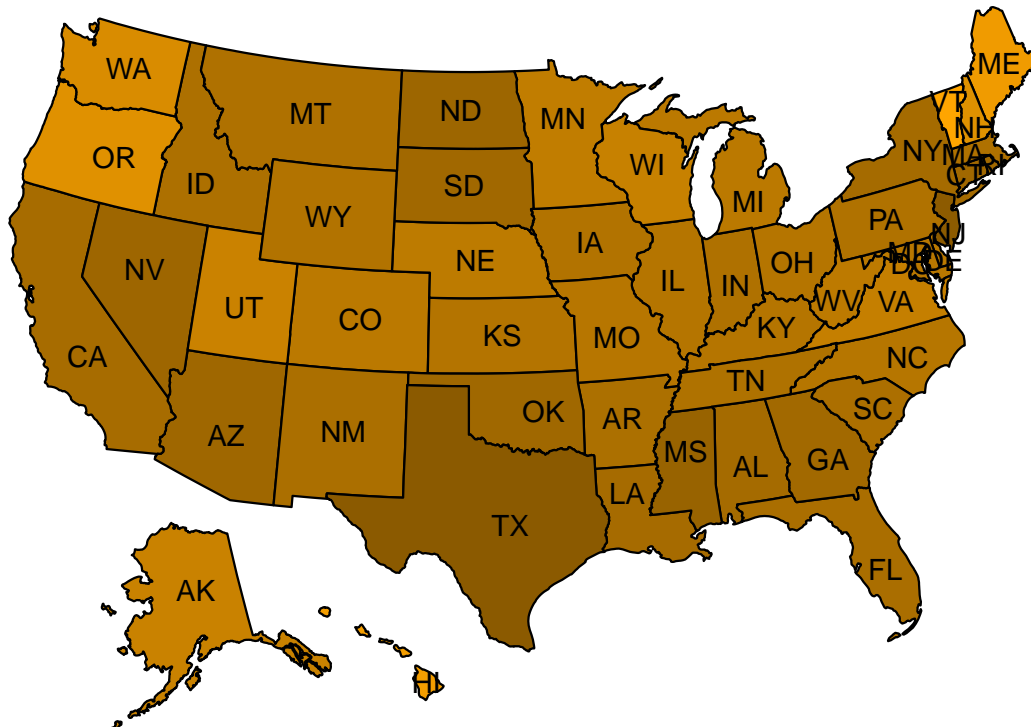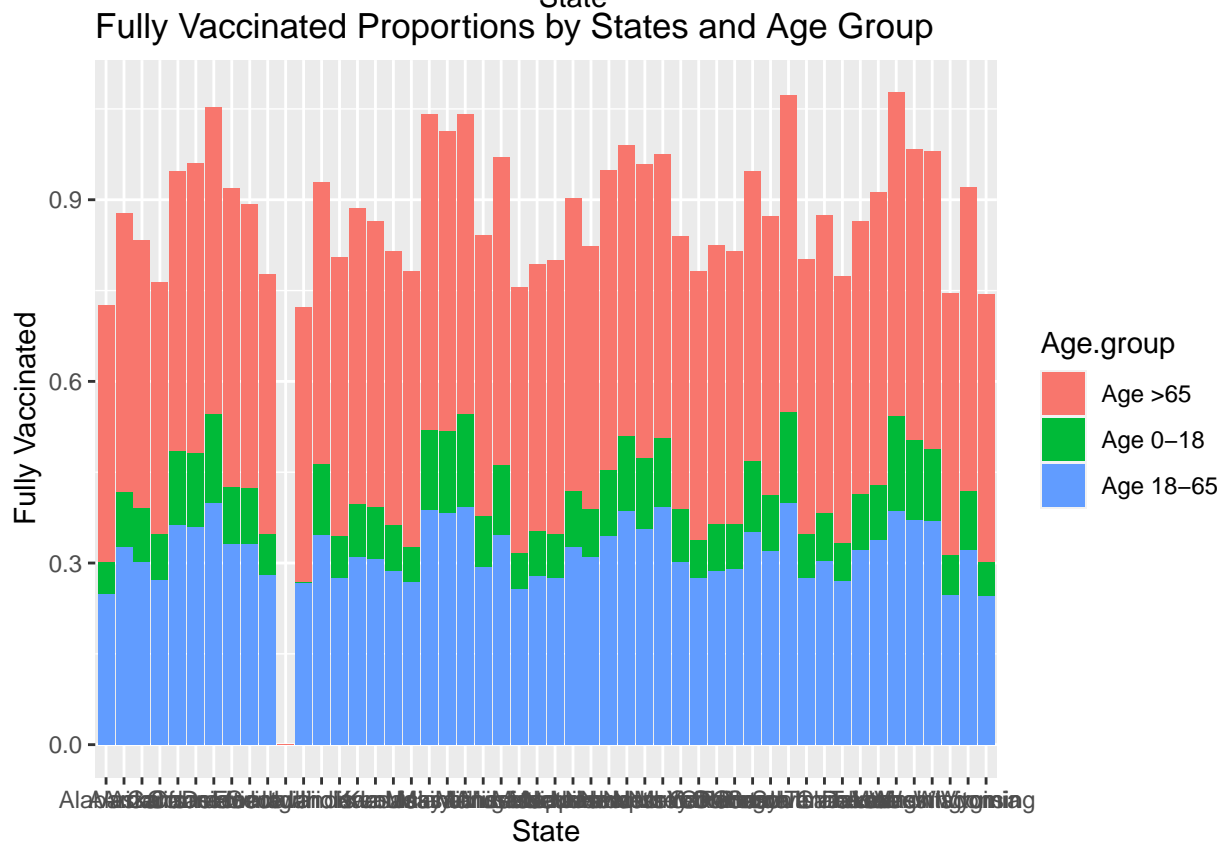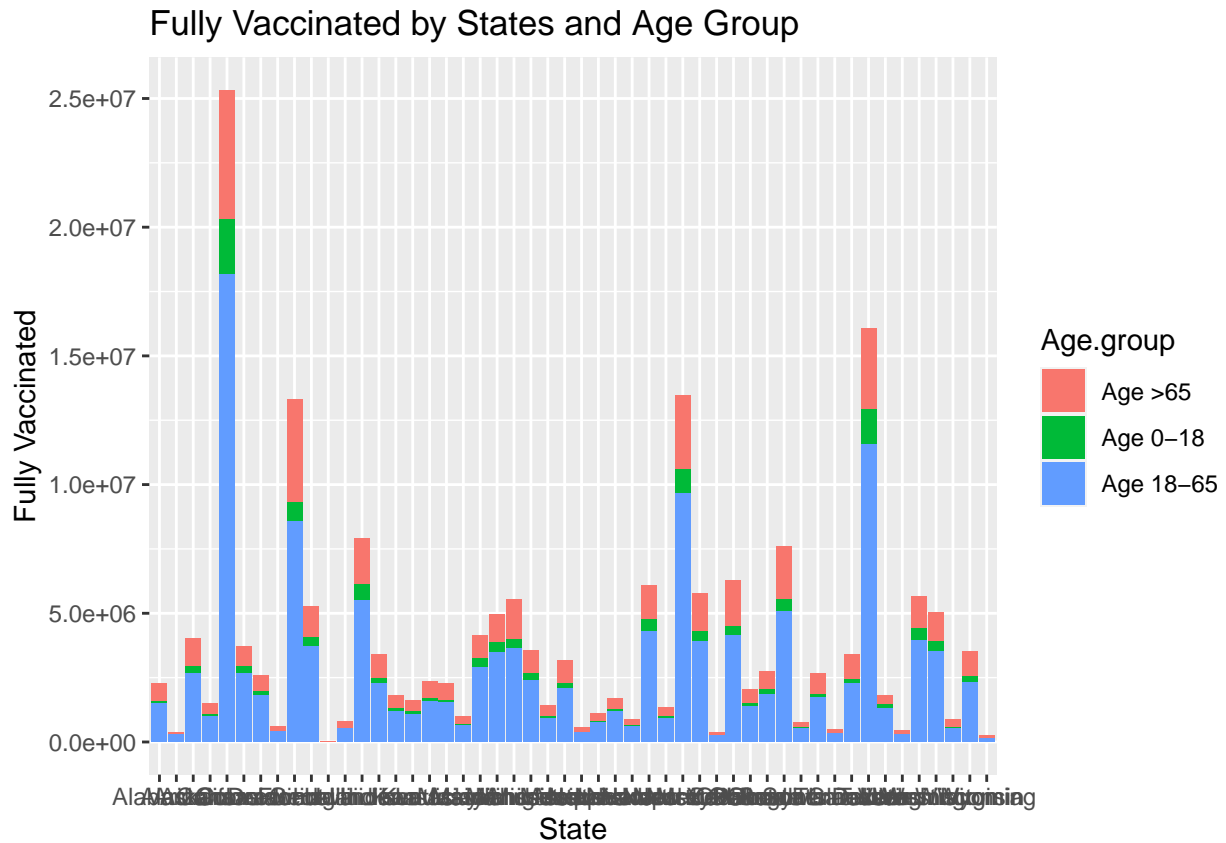
# References

# Figures

## COVID−19 Deaths by States and Age Group



## COVID−19 Deaths Proportions by States and Age Group
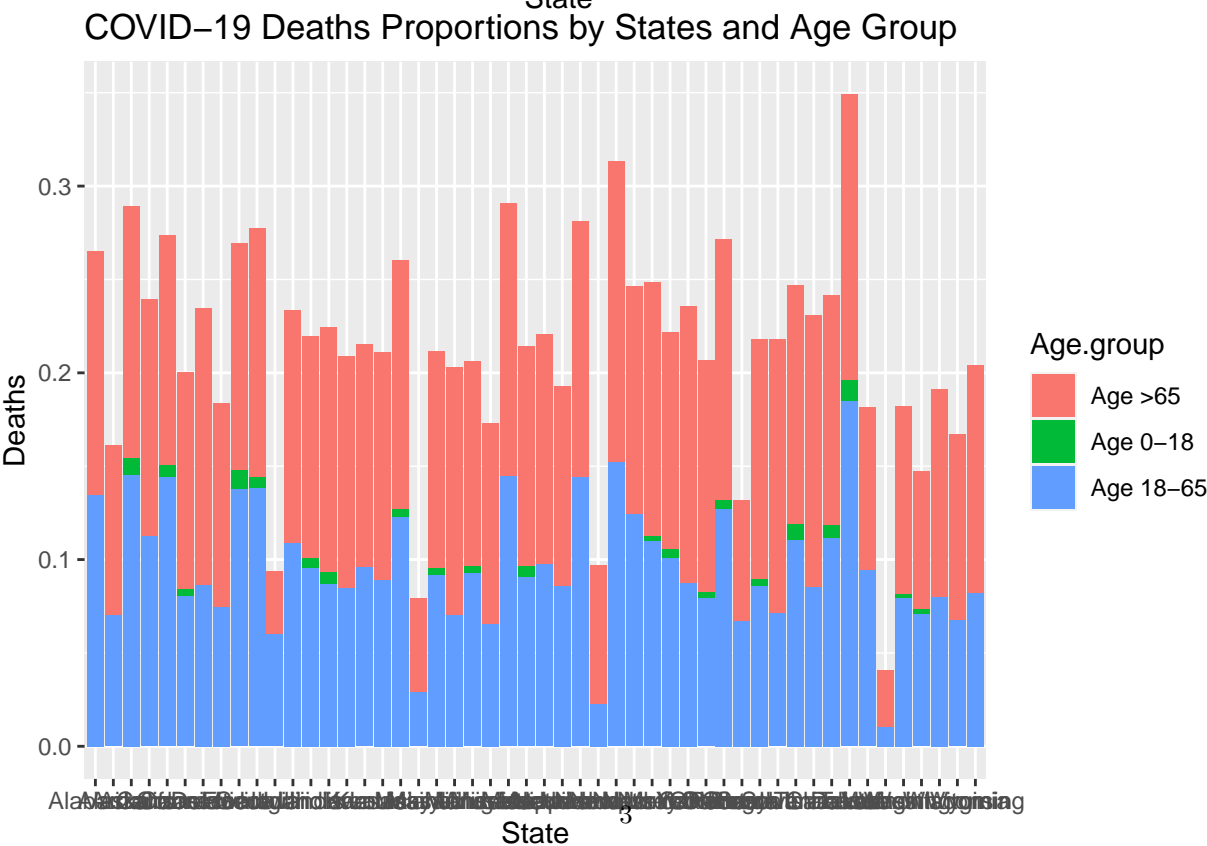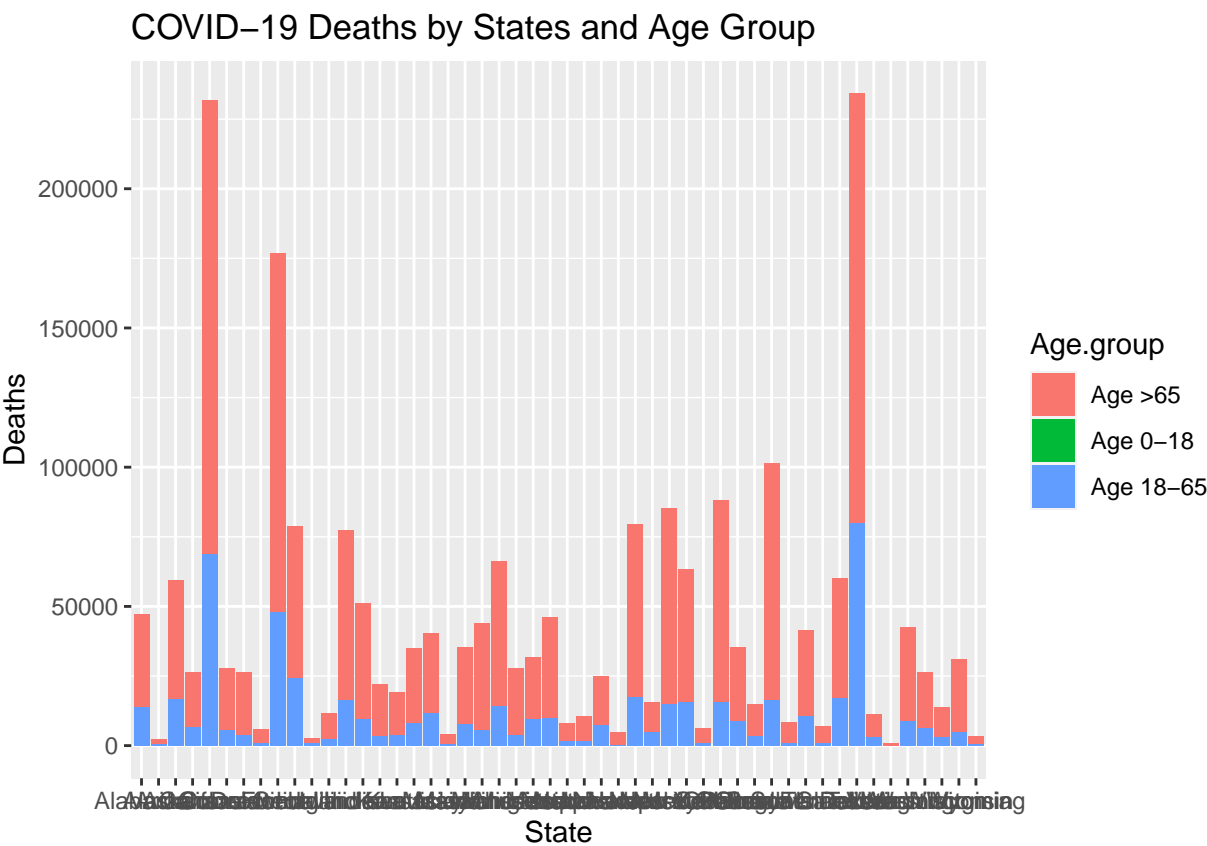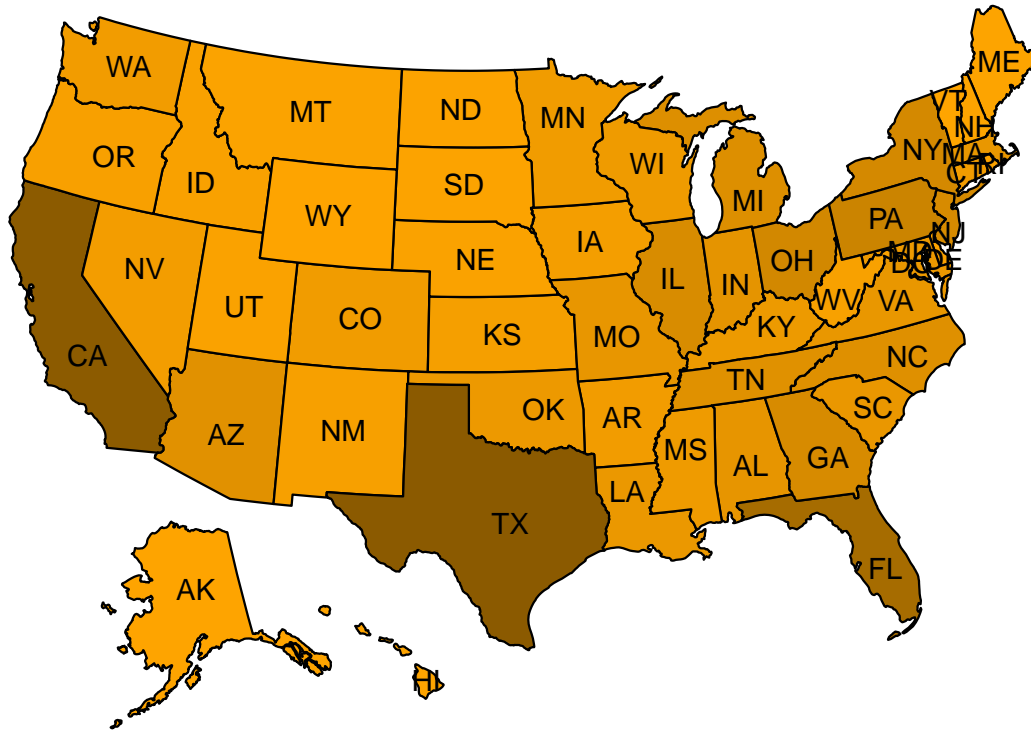
COVID−19 Deaths



Proportion of COVID−19 Deaths

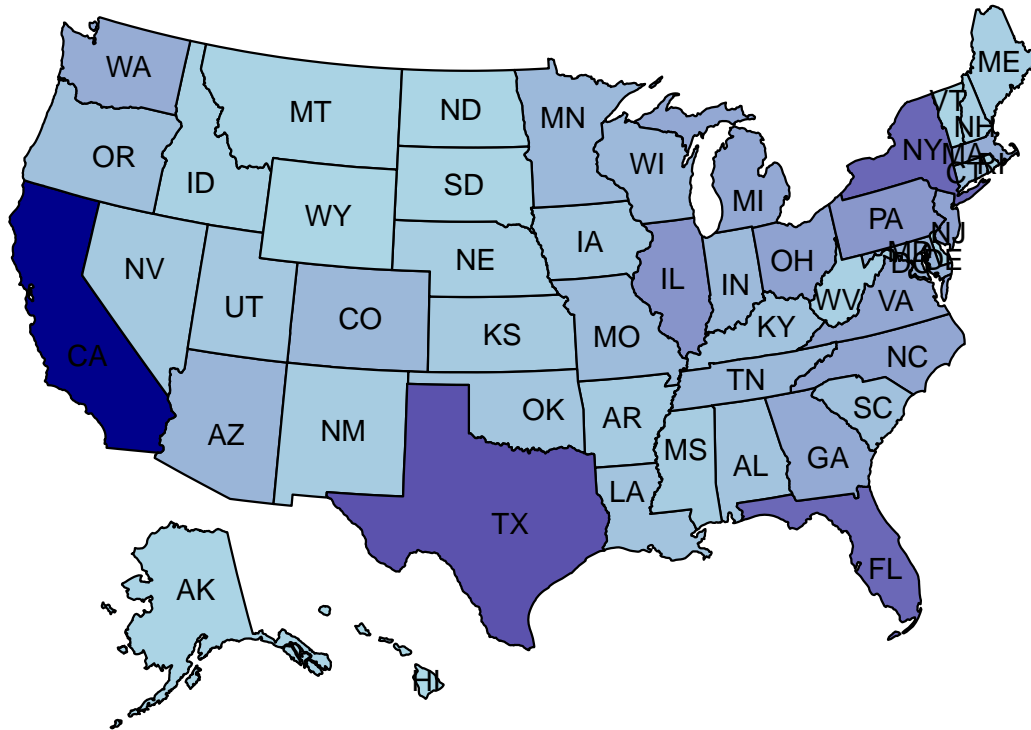Fully Vaccinated by States and Age Group



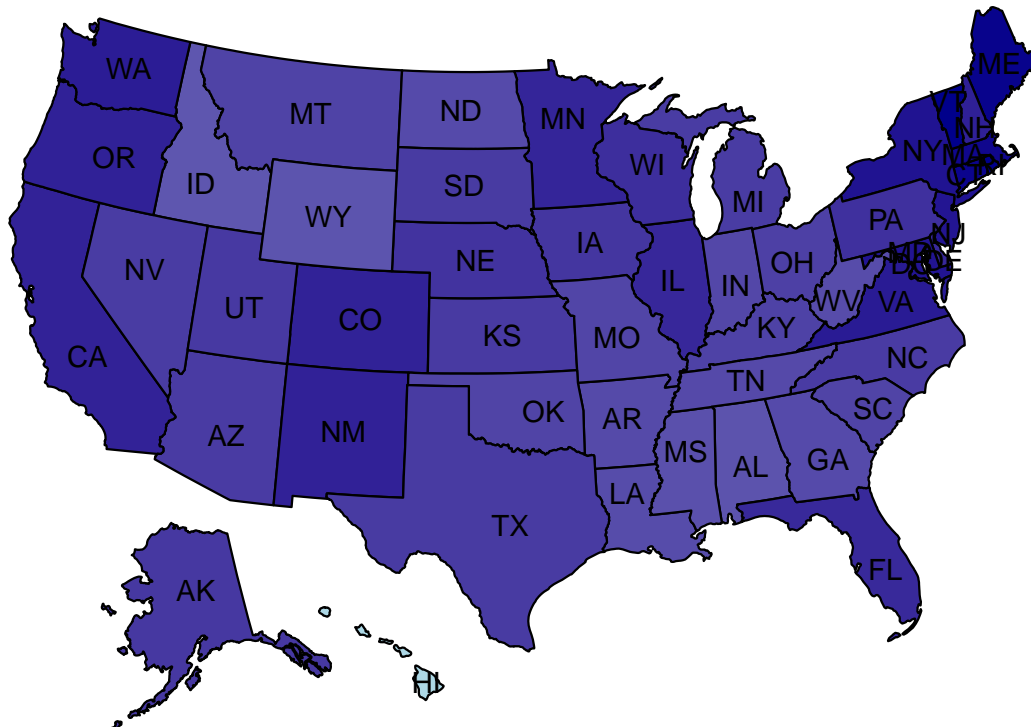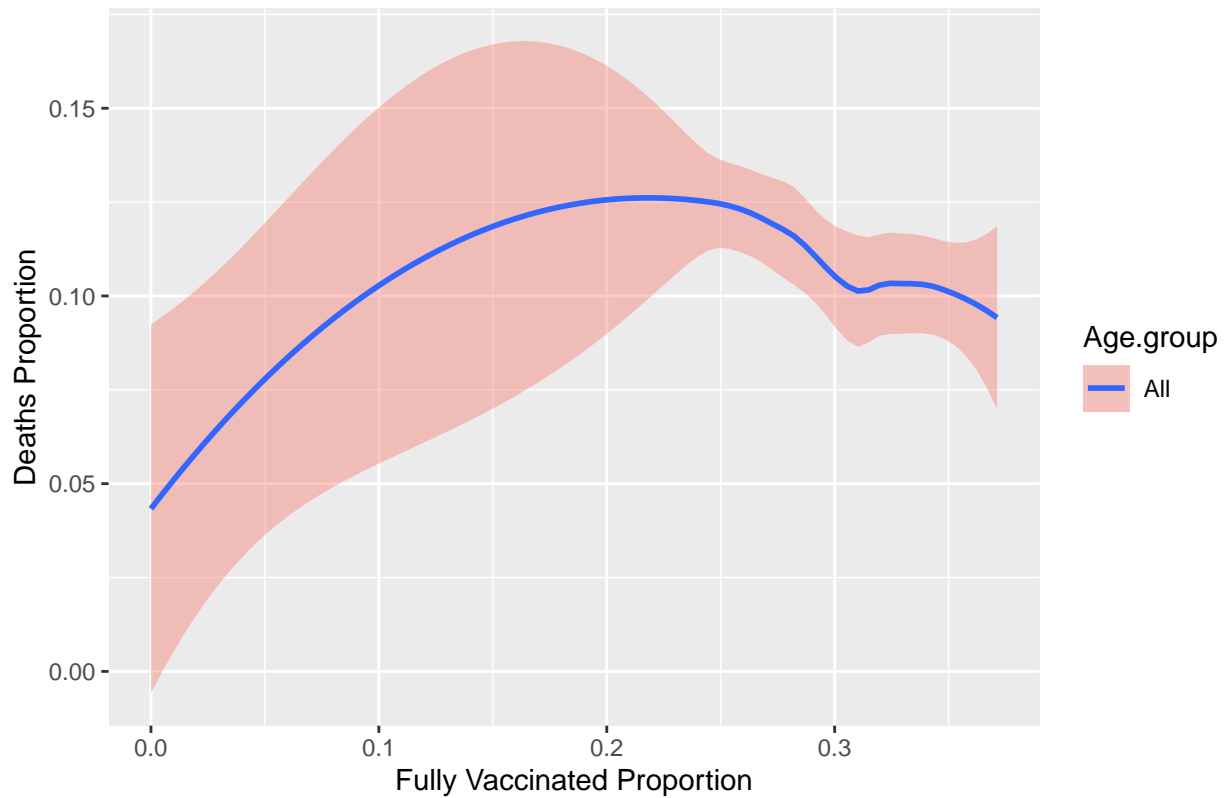Fully Vaccinated Proportions by States and Age Group

Fully Vaccinated People



Proportion of Fully Vaccinated People

Fully Vaccinated Proportion vs. Deaths Proportion



Fully Vaccinated Proportion vs. Deaths Proportion