Code

Texas Internal Migration

Comparing ACS data for 2019 and 2021

My-An Pham

14 March, 2023

0.1 Objectives of Project

In looking for interesting topics to research, I found that Forbes Home (https://www.forbes.com/home-improvement/features/states-move-to-from/) listed Texas as the #1 State that people are moving to in the United States (U.S.) based on the United States Postal Service's (USPS) address-change-request data. The article also states that Texas gained 12,700 residents in 2021. This peaked my interest in Texas's migration patterns, which is why I am investigating the following questions:

How do 2021 migration rates compare to that of a few years prior (2019) in relation to Texas? - Out of curiosity for migration rates regarding Texas being an increasing trend that has been developing for a few years or if it was a recent and new change.

Are there any patterns or differences in where people are moving between the years 2021 and 2019? - As a more elaborate exploration of the previous question, we can understand patterns and differences between U.S. migration in relation to Texas to draw more conclusions on why migration is happening.

Does the American Community Survey (ACS) depict similar data to USPS' claims about Texas being the #1 state people are moving to? - This will help us fact check claims made by USPS, seeing if different methods of collecting data affect results.

0.2 Data and methods

In order to answer the previous questions, I have gathered data from the American Community Survey (ACS) and collected information from the years needed:

```
-   2019 ACS Data
-   2021 ACS Data
```

Using this data, I will compare both years' inmigration, outmigration, netmigration and migration efficiency for each state in relation to Texas. I will also calculate each state's overall netmigration to compare to USPS' data. I will be using the PERWT values from the ACS data to generate these comparisons, which is the amount of people that a given person in the sample represents.

0.3 Load Data and Create Base Datasets/Variables

In the following code chunk, we will load in the packages needed to handle our data and generate our maps. We will also download, extract, and clean our data as well as download a shapefile of the U.S. so that we can easily manipulate, visualize, and analyze our data.

Code

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✔ ggplot2 3.4.0      ✔ purrr   0.3.5
## ✔ tibble  3.1.8      ✔ dplyr   1.0.10
## ✔ tidyr   1.2.1      ✔ stringr 1.5.0
## ✔ readr   2.1.3      ✔ forcats 0.5.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

Code

```
## Linking to GEOS 3.9.3, GDAL 3.5.2, PROJ 8.2.1; sf_use_s2() is TRUE
```

Code

```
## `summarise()` has grouped output by 'PREV_STATENAME', 'STATENAME'. You can
## override using the `.groups` argument.
```

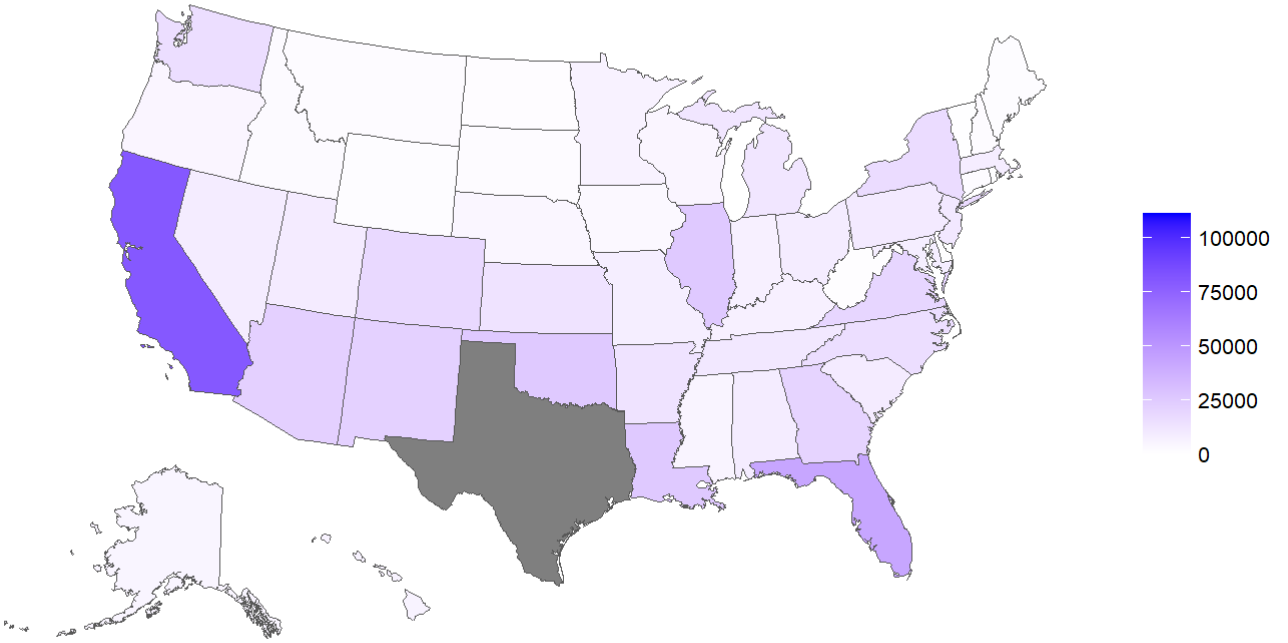
0.4 Compare State Inmigration Relative to Texas

First, we will look at inmigration relative to Texas, which is the number of people who are moving into Texas from each state. The following code first filters the 2019 ACS data to focus on Texas and the amount of people who have left their state to move to Texas. Then, using the U.S. shapefile data and ggplot in R we display the inmigration rates on a map of the United States, with higher rates being more blue and lower rates being more white. Then, we repeat these steps for the 2021 ACS data. The limits are the same for both maps and were determined by the highest and lowest values between both years to ease recognition of patterns and differences between the two years.

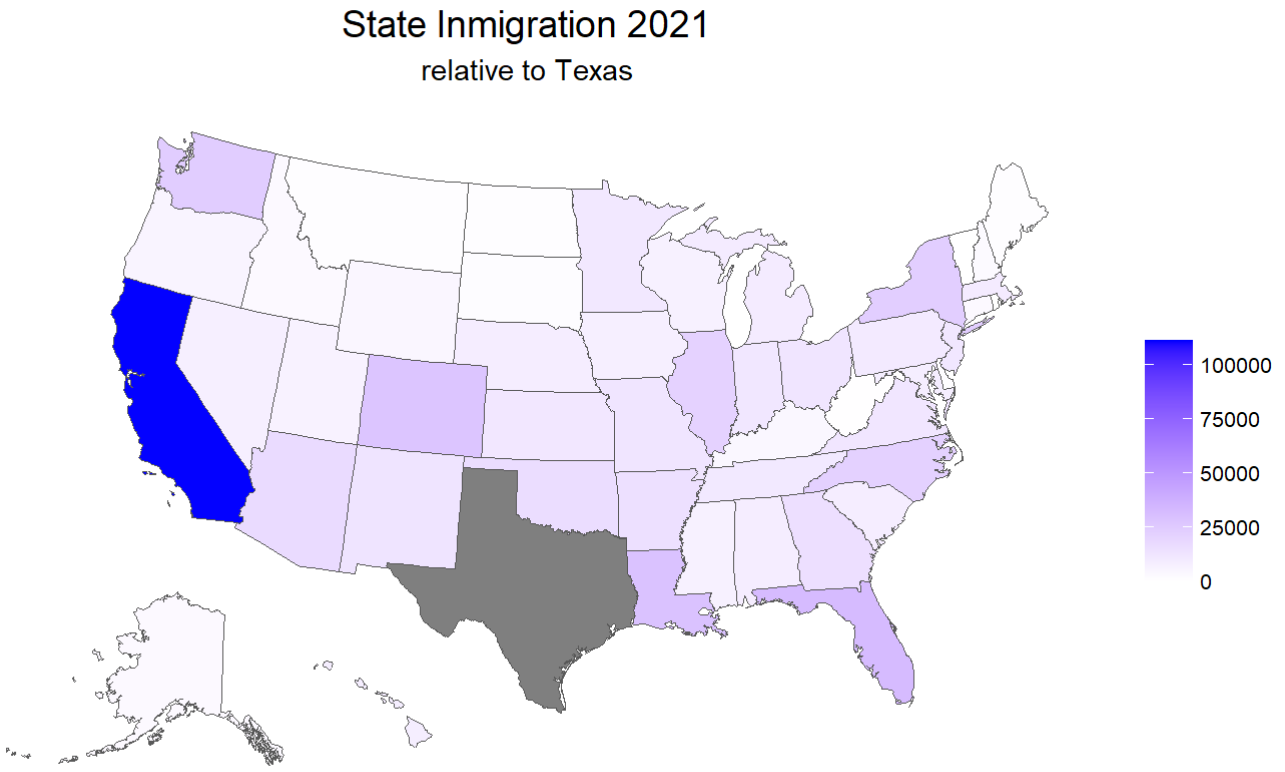
Code

State Inmigration 2019

relative to Texas



Code



In looking at the two maps produced, we can see that California has the highest rate of inmigration into Texas in both maps but is much higher in 2021 (signified by dark blue). For the most part, it seems as though in the southern half of the United States for the 2019 map, most of the states are ever so slighlty more purple than in the 2021 map, but in the northern half most of the states are ever so slightly less purple than in the 2021 map. Visually, the 2021 map seems to have more spread of purple through the whole map, with many being a very light purple and the 2019 map seems to have less states that are purple but has a few more that are a moderately light purple. There are many states that seem unchanged between the two years and stayed more on the white side, like Oregon, Idaho, Montana, North Dakota, South Dakota, West Virginia, Nevada, Mississippi, Vermont, New Hampshire, Connecticut, Rhode Island, Maine, and Delaware, meaning that between both years they had low rates of inmigration into Texas.

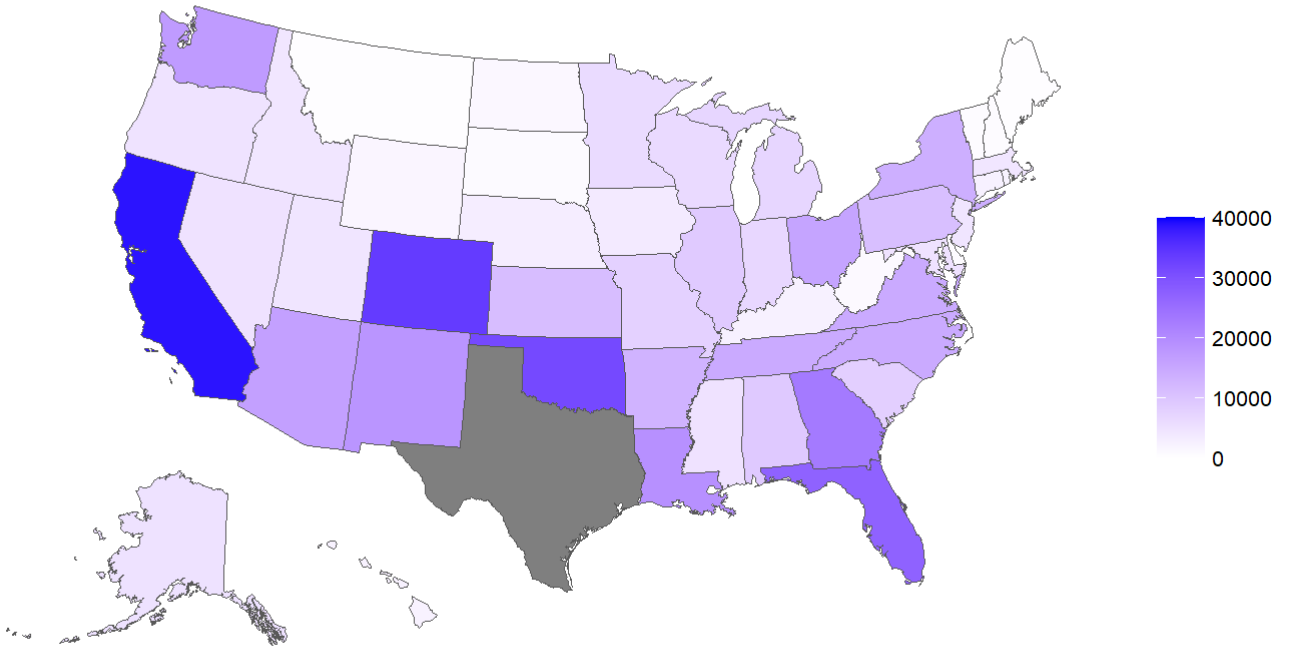
0.5 Compare State Outmigration Relative to Texas

Now we are going to be looking at outmigration in relation to Texas, which is the number of people who are moving out of Texas to other states. The code first filters the 2019 ACS data to samples with samples that were previously in Texas but are currently in a different state. Next, using the U.S. shapefile data and ggplot in R, the outmigration rates are put on a U.S. map. Then the steps are repeated for 2021 ACS data. The limits are determined by the highest and lowest outmigration values between both years, so that it is easier to visually compare the two maps. The more blue a state is the higher the outmigration rate and the more white a state is the lower the outmigration rate.

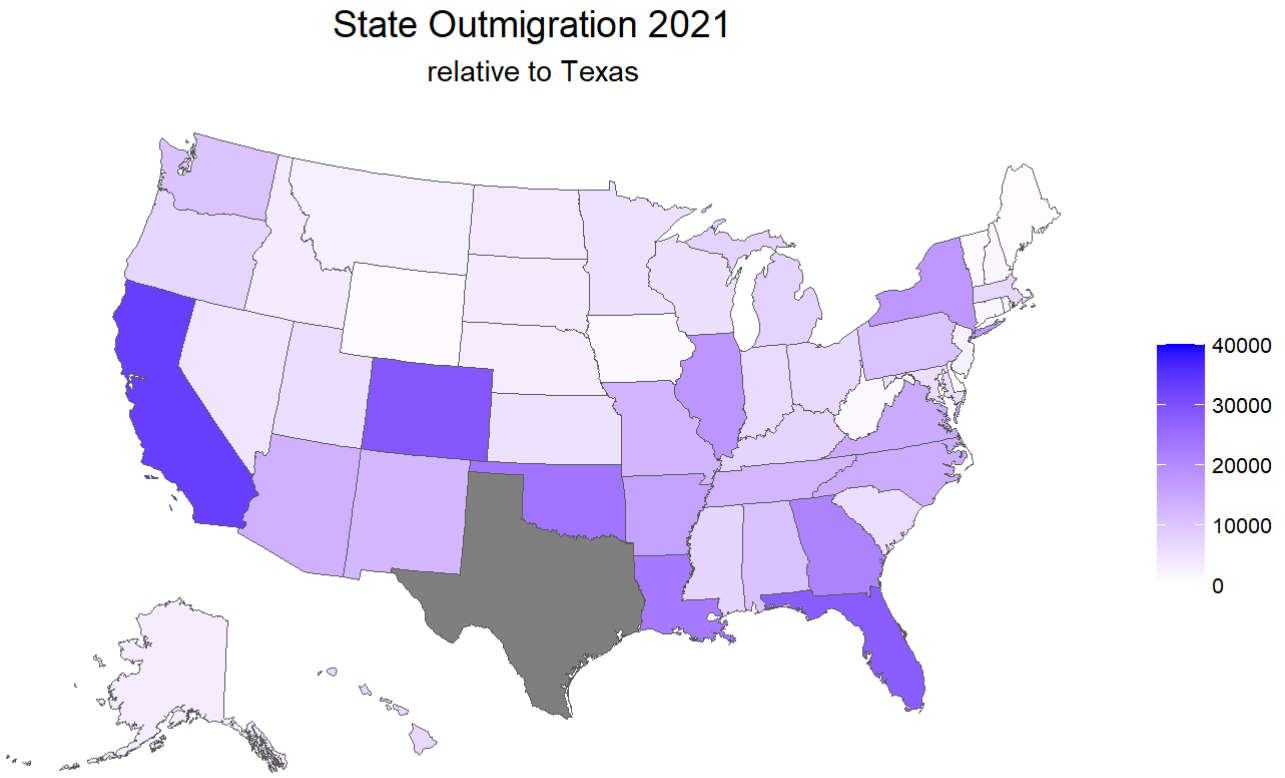
Code

State Outmigration 2019

relative to Texas



Code



It is important to note that these outmigration maps compared to the inmigration maps generated previously may seem like they have higher rates due to more states being dark purple and blue, but this assumption is not very accurate due to the limits of the maps being different (the hightest value on the inmigration maps are much higher than these outmigration maps). In the 2019 map, California, Colorado, Nebraska, and Florida have the highest outmigrations out of Texas (signified by dark purples and blues). We can see that in 2021 these states are still the states with the highest outmigration rates in 2021, but they have lower rates than the 2019 map with the exception of Florida, which stays the same. Some states are a moderate light purple in 2019 and decrease to a lighter purple in 2021 (meaning their outmigration decreased), like Arizona, New Mexico, Tennessee, Ohio, and Washington. A few states change from a moderate light purple to a darker purple from 2019 to 2021 (meaning their outmigration increased), like Georgia, New York, Illinois, and Louisianna. 2021 seems to have more states overall that are not white, but some states stay white in both maps (meaning the outmigration is zero or very low in both maps), including West Virginia, Wyoming, Vermont, Maine, Connecticut, Rhode Island, and Delaware.

0.6 Compare State Netmigration Relative to Texas

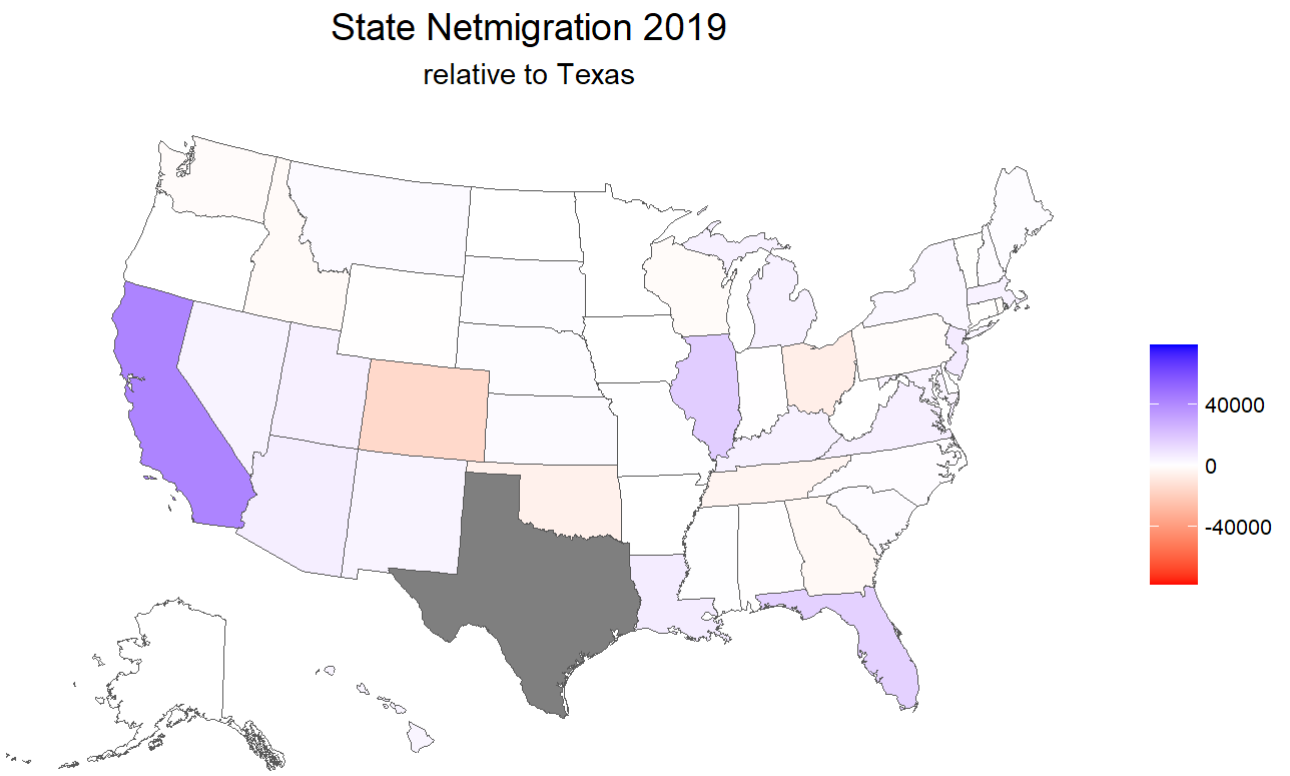
Since the outmigration and inmigration maps have different limits, comparing the two as they are now is rather difficult. Because of this, we are going to be calculating the Netmigration for both years in relation to Texas for each state, which is done with the equation below:

Netmigration = Inmigration - Outmigration

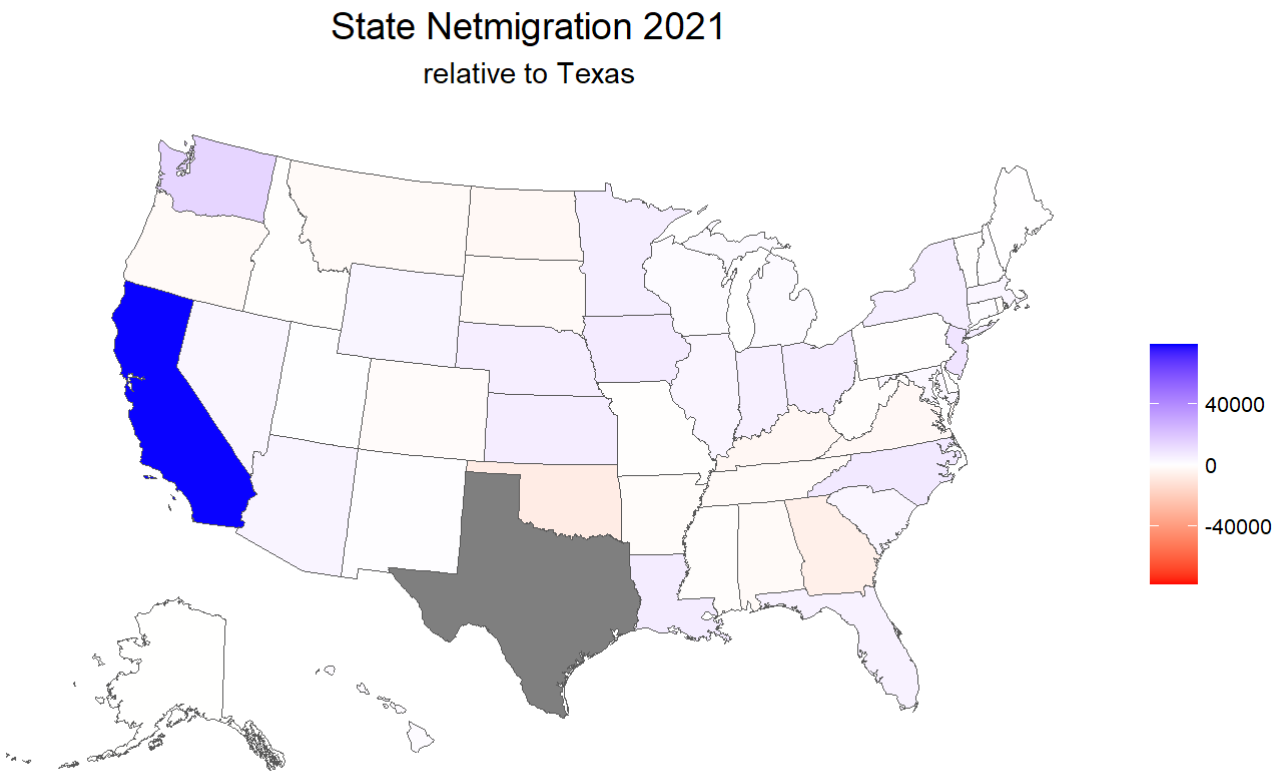
This means that a positive netmigration indicates more inmigration than outmigration (which on our map will be more purple), a negative netmigration indicates more outmigration than inmigration (which on our map will be more red), and a netmigration closer to zero indicates that the inmigration and outmigration are closer to being equal to each other (which on our map will be more white).

To do this, we combine the 2019 Texas inmigration and outmigration data into one dataframe, making sure to only focus on data relating to Texas. After, we create a new column for netmigration that subtracts the Texas inmigration by outmigration for each state. We then use R's ggplot and the U.S. shapefile data to map the data. These steps are repeated for the 2021 ACS data and the limits of both maps are the same and based on the netmigration value farthest from zero between the two datasets so that we can compare the two maps with ease.

Code



Code



Overall, we can see that the 2021 map has more states that are purple but many of the states are a very light purple, meaning that there is more inmigration into Texas the outmigration out of Texas for these states, but not by much. This is due to many of the white states in 2019 (indicating little to no difference in inmigration and outmigration rates) becoming light purple in 2021. Many states stay white between both maps: Alaska, Hawaii, Louisianna, Mississipi, West Virginia, Pennisylvania, Vermont, New Hampshire, Maine, Connecticut, Delaware, Arkansas. In 2019, Illinois and Florida are moderately light purple so there is more inmigration than outmigration, but this difference decreases in 2021 as both states become a lighter purple. Also in 2019, out of all states with a higher outmigration than inmigration (Colorado, Nebraska, Tennessee, Georgia, Ohio), Colorado is the most pink meaning it has the largest difference. Each of these states generally either become more white or become purple in 2021, meaning that the difference between inmigration and outmigration decreases for all of them with some changing to have a higher inmigration than outmigration. Montana, North Dakota, and South Dakota change from white to a very light pink between 2019 and 2021, meaning they all had inmigrations and outmigrations that were about equal but change to have slightly more outmigration. California's netmigration indicates a higher inmigration for both years, but the difference increases immensely from 2019 to 2021 as the purple turns into a dark blue. Something interesting to note, is that in 2021 Washington has the second highest positive netmigration behind California, but it is still significantly lower than California's (as it is a darker purple than the other states but is still a moderately light purple).

0.7 Compare Migration Efficiency

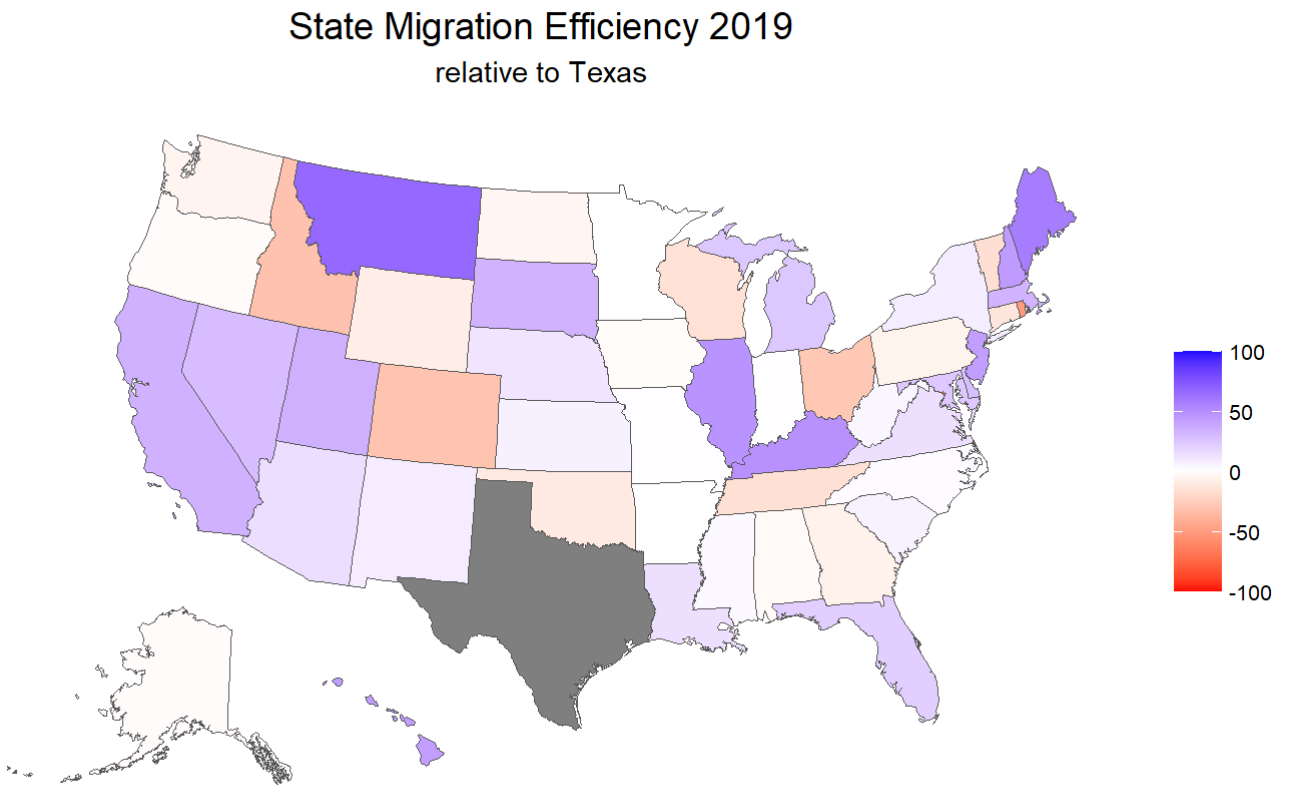
Sometimes, it might be more beneficial to calculate and analyze a standardization of netmigration, as it would allow us to control the effects of population size. This can be done using migration efficiency, which is found using the following equation:

Migration Efficiency = 100 * (netmigration/(inmigration + outmigration))

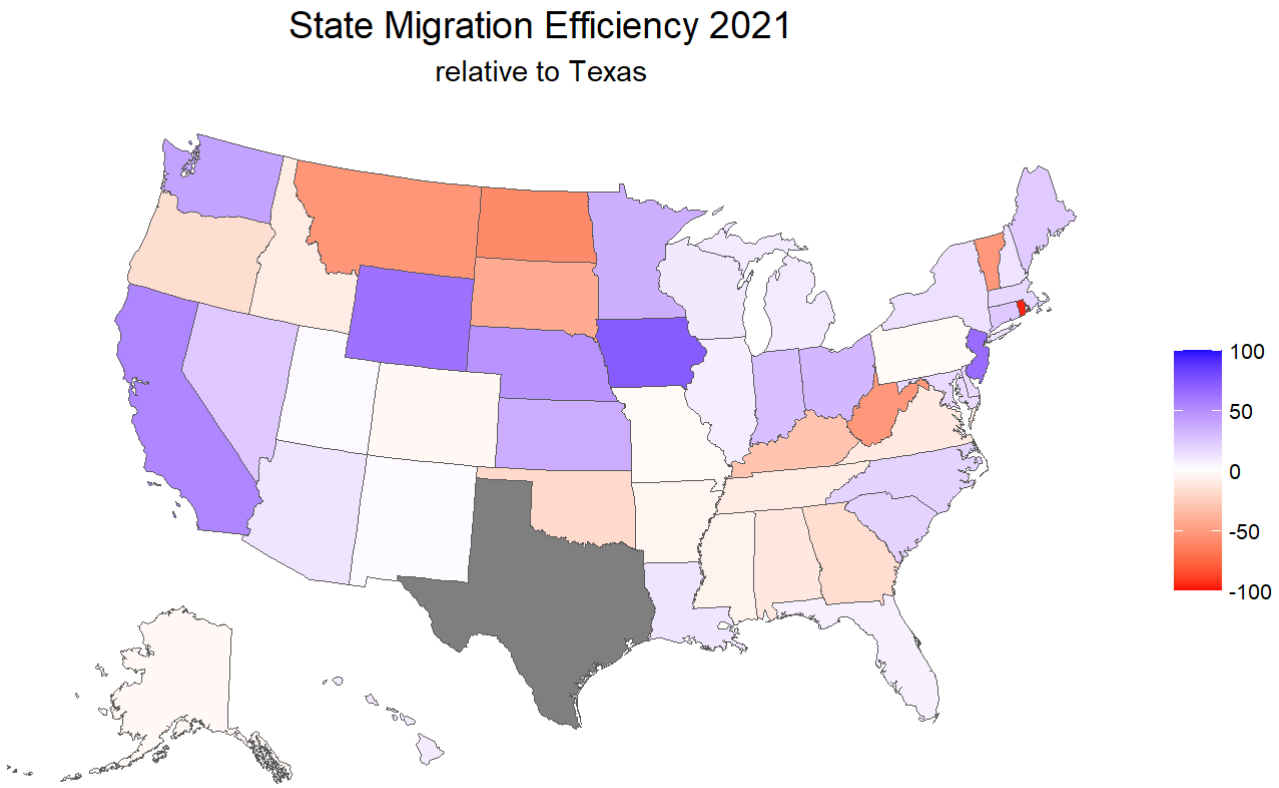
To make this easier to understand, we can break its meaning into two parts. The first, is that similar to netmigration a positive migration efficiency indicates: inmigration (into Texas from a given state) is higher than its outmigration (out of Texas into the given state); a negative migration efficiency indicates: outmigration is higher than inmigration. The second, is that the father the migration efficiency is from zero, the closer the netmigration is to the total flow of inmigration and outmigration. To put this together, a high positive migration efficiency indicates a high percent of netmigration to total flow from a high inmigration and a high negative migration efficiency indicates a high percent of netmigration to total flow from a high outmigration.

In the code below, we take the 2019 ACS data and use our previous dataset that includes the netmigration, inmigration, and outmigration, to create a column that calculates the migration efficiency for each state in relation to Texas. Then, using ggplot and the U.S. shapefile data, we create a map of the migration efficiencies in each state. These steps are repeated for the 2021 data and the limits for both maps are set to be -100 and 100 for both, since migration efficiencies are given in values out of 100 (making it easier to compare efficiencies within a single map and efficiencies between the two maps).

Code



Code



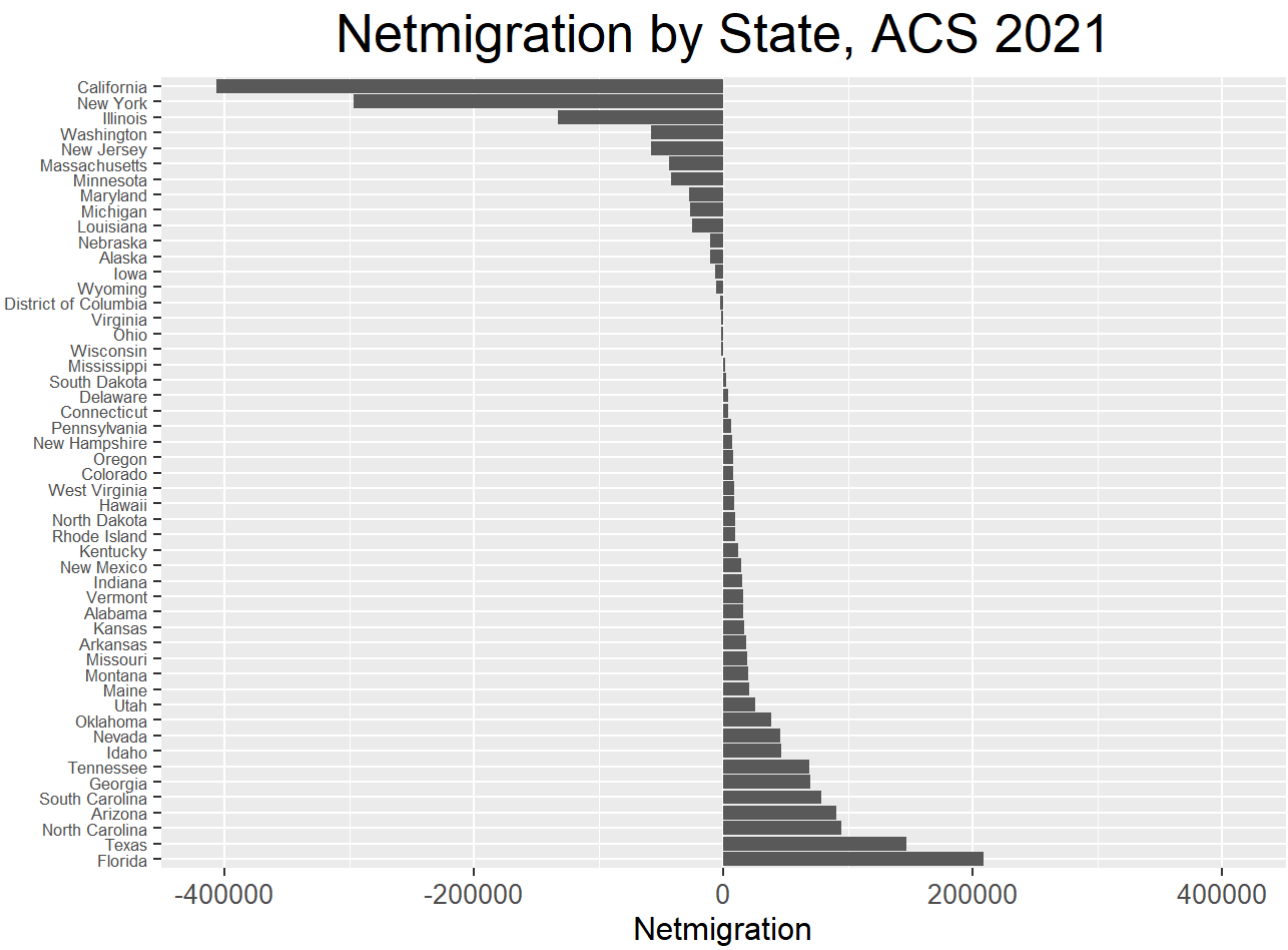
In only a matter of years, we see a drastic change in the migration efficiencies between 2019 and 2021. The only states that do not seemingly observe any relatively drastic change, are Alaska, Missouri, New York, Maryland, and Delaware. All other states either distinctly change shades of their original color, or completely change from positive to negative or negative to positive (signified by its change in red/pink or purple/blue). There are also clearly more states with higher negative migration efficiencies in 2021 than in 2019 (signified by darker reds). Because of this, it seems as though in 2019 there are more higher positive migration efficiencies than higher negative migration efficiencies, but in 2021 the amount seems more balanced.

0.8 Compare 2021 State Netmigration Overall

Since the Forbes Home (https://www.forbes.com/home-improvement/features/states-move-to-from/) article used the terminology “gained residents”, we are making the assumption that they drew conclusions based on netmigration. To compare these conclusions to ACS data, we must look at each state’s netmigration, using the same netmigration equation above: subtracting overall inflow by overall outflow. This will allow us to see which state has gained the most residents (in other words, has the highest netmigration due to inmigration or highest positive netmigration).

In the code below, we generated that takes the 2021 ACS data and calculates the total netmigration for each state. Then using ggplot, a ranked plot of netmigration by state is created.

Code



In looking at this plot, we can see that the ACS 2021 data differs from the USPS data as Texas is the second state to gain the most people as opposed to the first. I would argue that in terms of Texas' ranking the results are similar considering first and second are close, given that there were fifty rankings that were given. Because of this, we can more confidently conclude that Texas had one of the highest net increases in migration for the year 2021. It is important to recognize however, the difference in gaining 150,000 residents with the ACS data to the 12,700 listed in the Forbes Home article is a much larger difference that would need further exploration. This leads us to ask, what causes these differences?

Upon further research, I found a NCOALink (https://postalpro.usps.com/mailing-and-shipping-services/NCOALink) information page that describes the framework of USPS' change-of-address data. Using this, are many inferences I can make about why there are differences between the ACS Data and the USPS data, but it is tough to say for sure considering the lack of clear citation in the Forbes Home article (I was only able to find a link to the USPS home page). Without being able to look at the USPS data myself, it is difficult to decipher where 12,700 comes from. Does this number include change-of-address for businesses? Does this number include individuals or just households? When they claim "gained residents" does that actually mean net gain like I assumed, or is it just total gain?

0.9 Conclusion

There are many interesting intricacies in comparing migration patterns of Texas between the years 2019 and 2021. Most interesting to me, is that only a matter of two years can generate differences in migration rates for Texas, with only a small handful of states having the same migration efficiency and the remaining changing rather drastically, as well as California staying consistent with having the highest positive

netmigration in both years, but seeing a large increase in netmigration comparing 2019 to 2021. As for how ACS data compares to USPS data, my inability to draw valid and accurate comparisons highlights the importance of data citation and struggles of finding accessible data.