

Matteo Ferniani, Kateryna Biriukova, Simone Garuffi

Lost in the Golden State

Exploring California's Depopulation Crisis

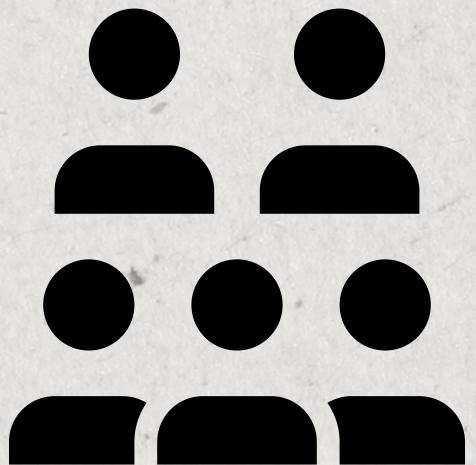


Problem and Audience



Problem

California is one of the richest states in the world (GDP per capita) with low unemployment and high income. In the last few years we have observed an unusual exodus of population, decrease in life standards, and unaffordable cost of living.

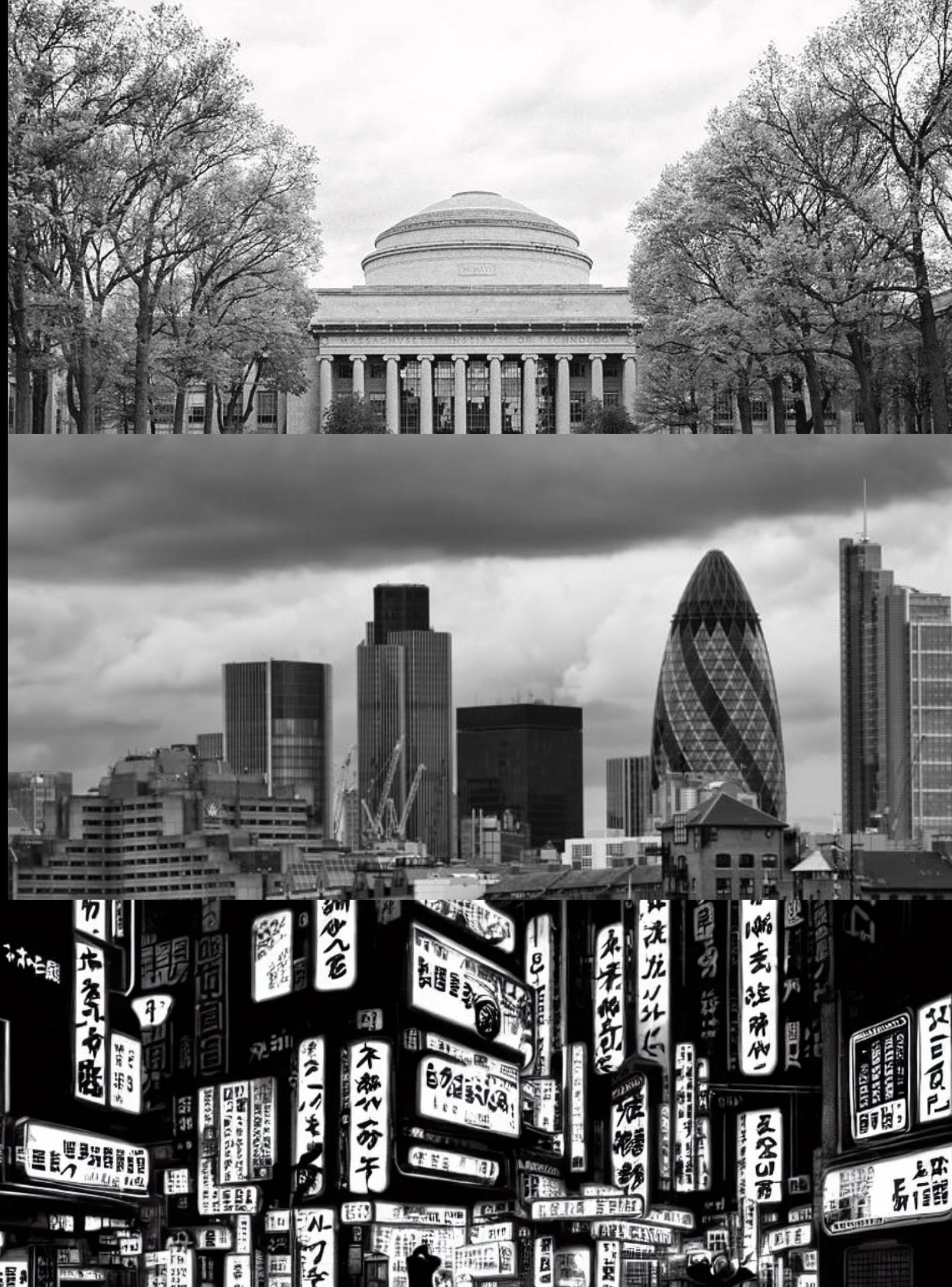


Audience

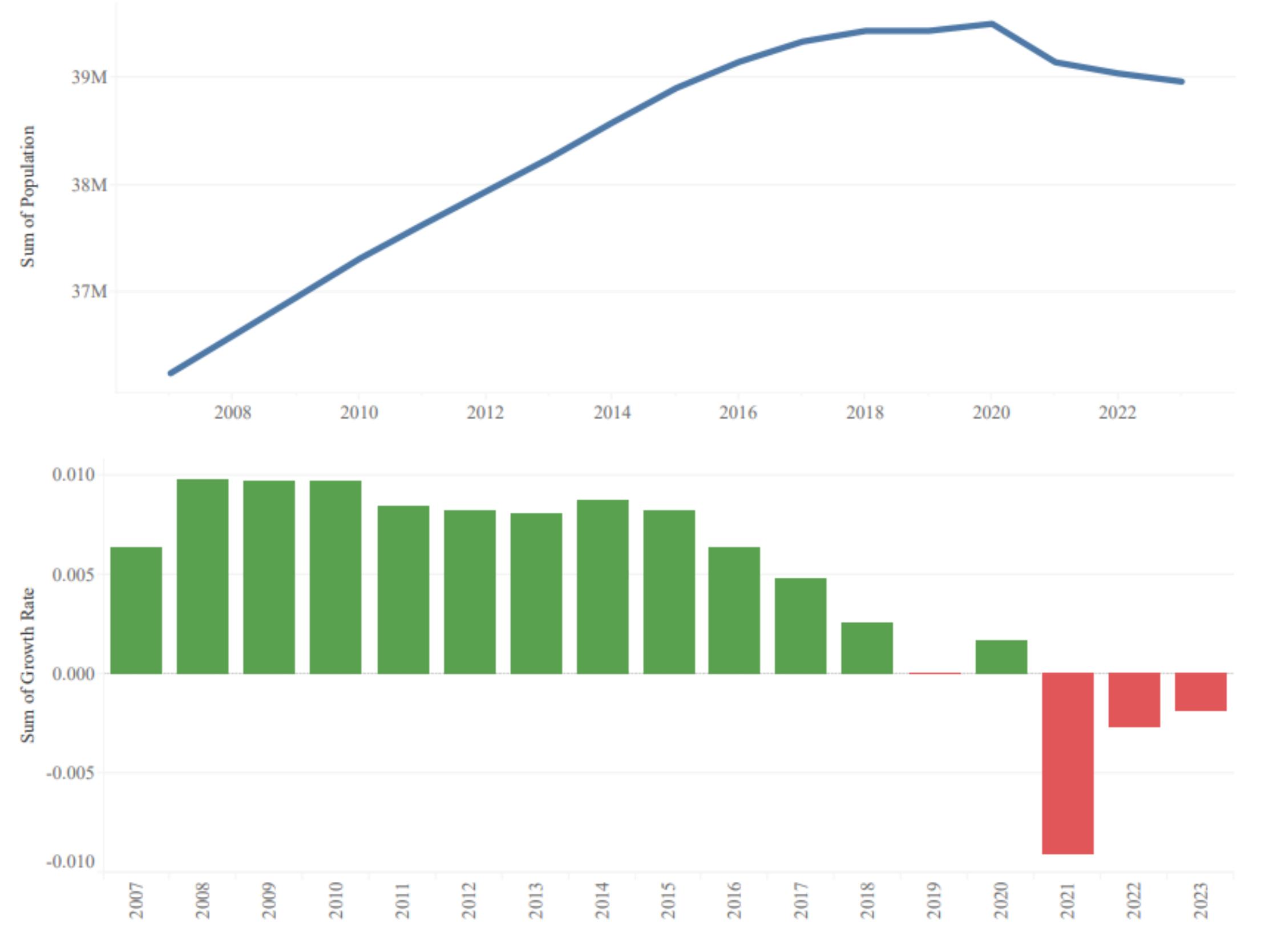
Governors of California is our main Audience, since they are potentially interested in the cause of the problem and possible ways to solve it.

Why is the chosen problem Important:

California, among the wealthiest states globally, faces critical issues despite its high GDP, abundant workforce, and robust welfare. Depopulation, soaring living costs, rising crime, homelessness, negative migration, and declining quality of life highlight urgent challenges. Studying California's plight offers vital lessons for cities like Boston, London, and Tokyo, poised to confront similar hurdles.



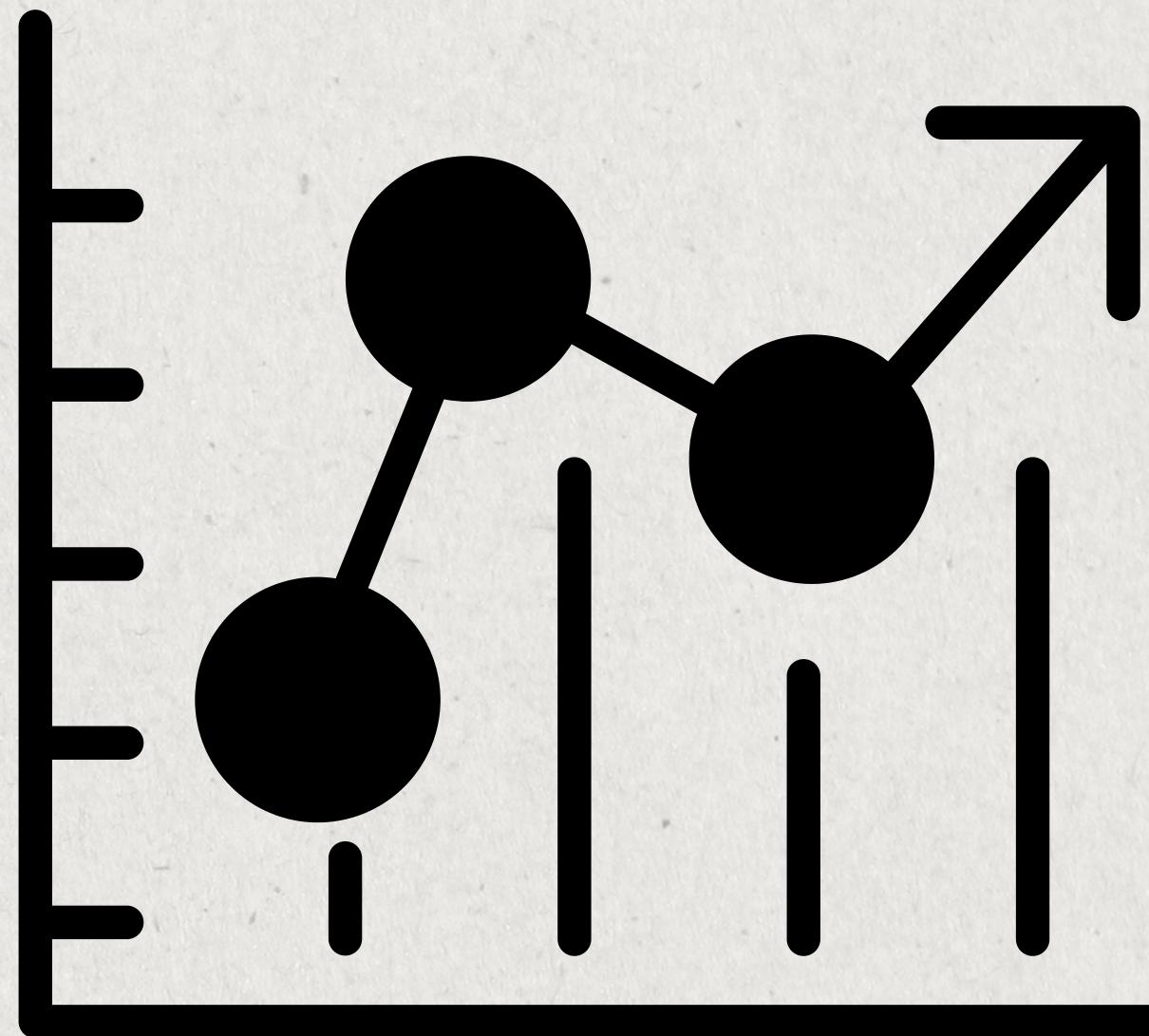
California Population and Growth Rate Change



How your analysis helps on understanding the problem

Data analysis can play a crucial role in addressing the challenges faced by California. By leveraging data analytics, policymakers can gain insights into the root causes of the population exodus, declining living standards, and soaring cost of living. Through thorough analysis of demographic trends, employment data, housing affordability metrics, and economic indicators, policymakers can identify areas that require targeted intervention.

Additionally, data-driven decision-making can facilitate the development and implementation of effective policies aimed at revitalizing the economy, improving infrastructure, expanding affordable housing options, and enhancing access to education and healthcare. By harnessing the power of data, California can strategize and prioritize initiatives to reverse negative trends and ensure sustainable growth and prosperity for its residents.



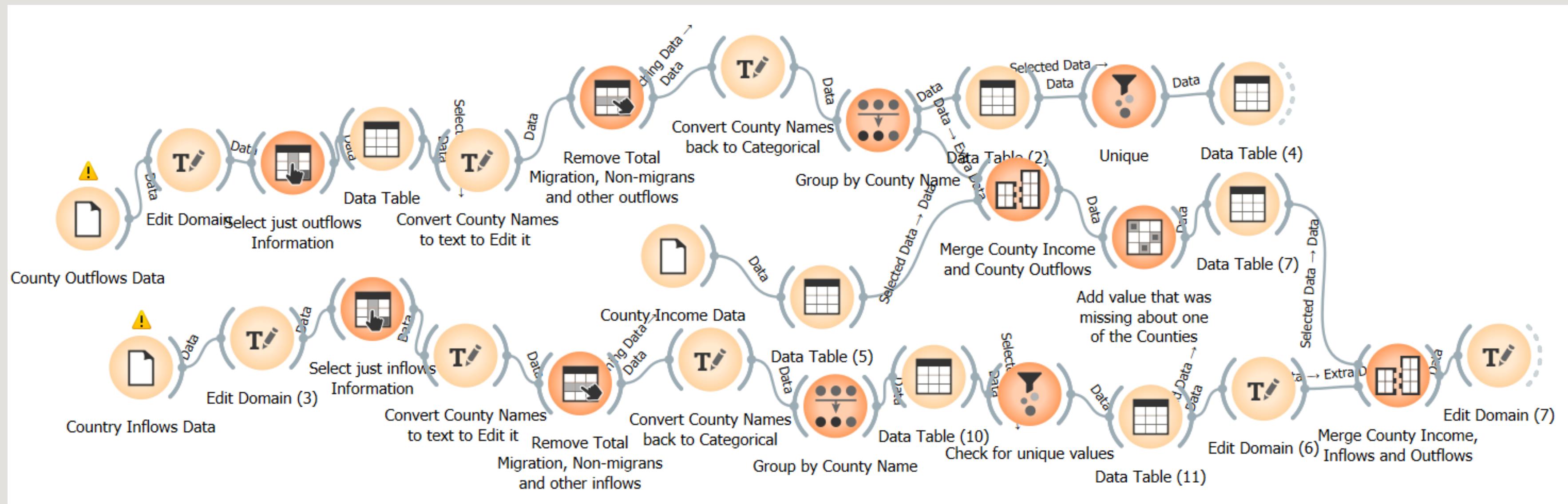
Data sources



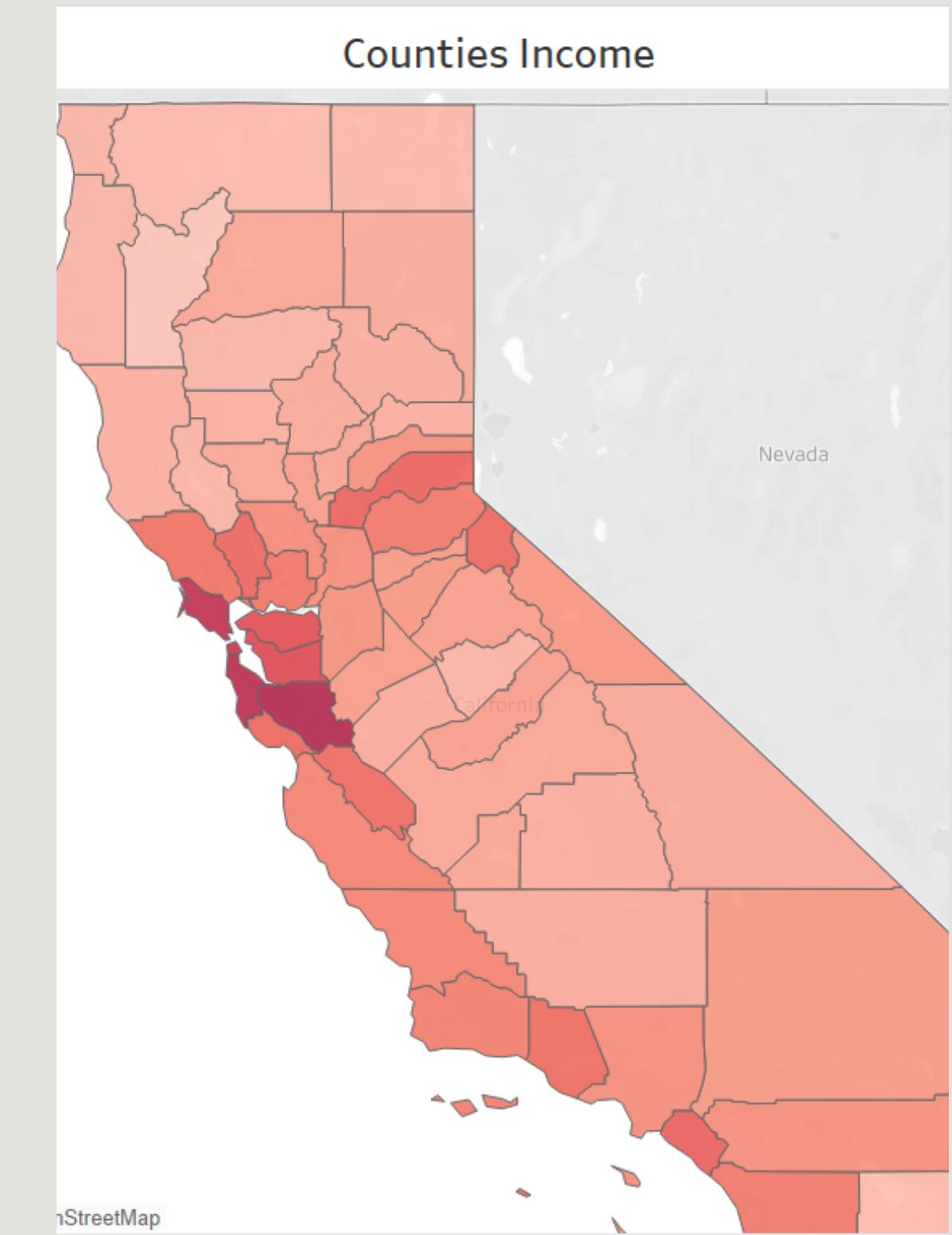
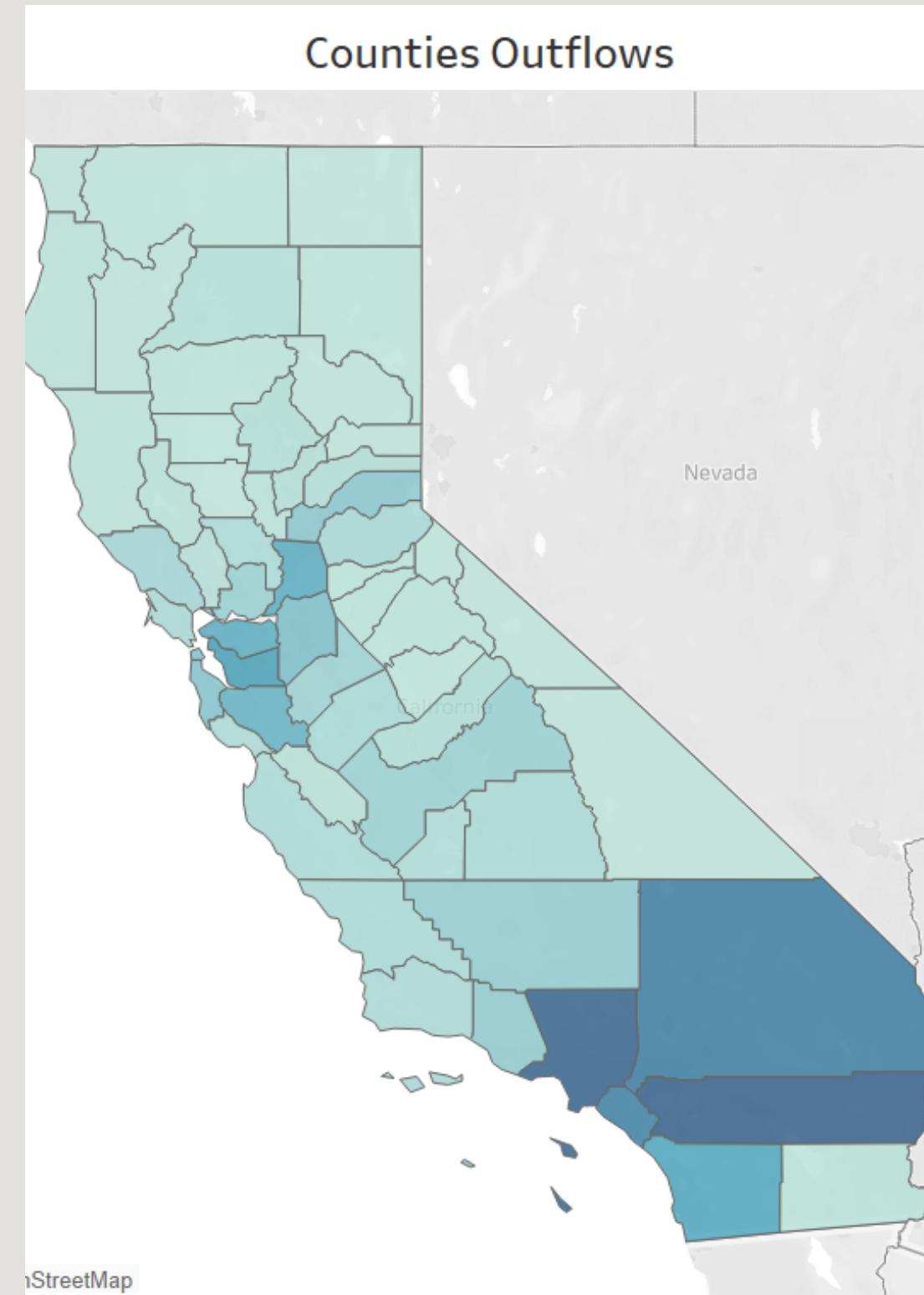
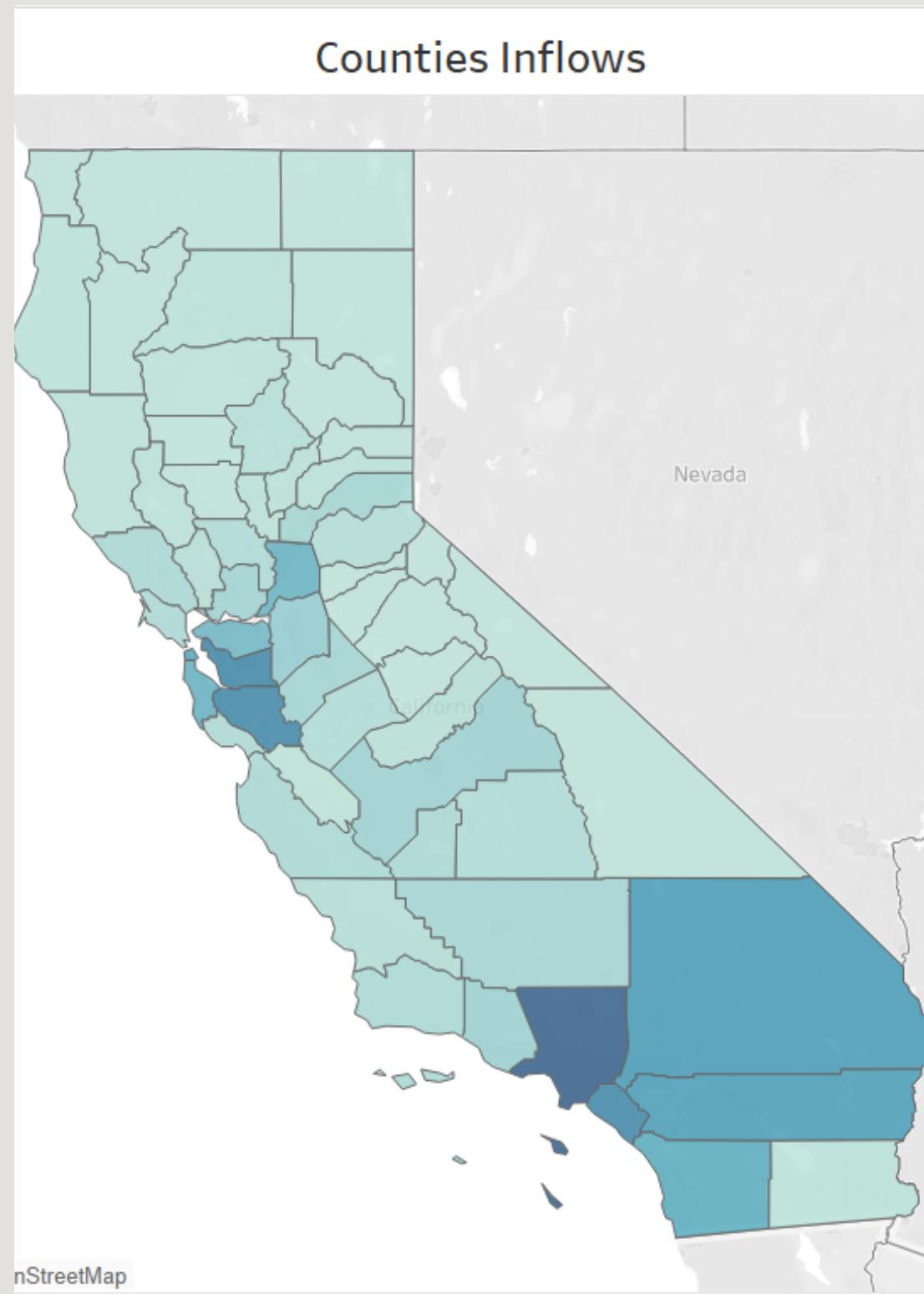
We used the following Data Sources: U.S. Census Bureau, Internal Revenue Service, Public Policy Institute of California, Statista

We got the following datasets: States' and Counties' inflows and outflows, Counties' income, California and Every State Population, Educated People migration data, Migration Per Income level, Work from Home Data, Gini Coefficient of California, Net Migration, US Metropolitan Area Cost of Living, Homeless population

Counties Information Orange Workflow:

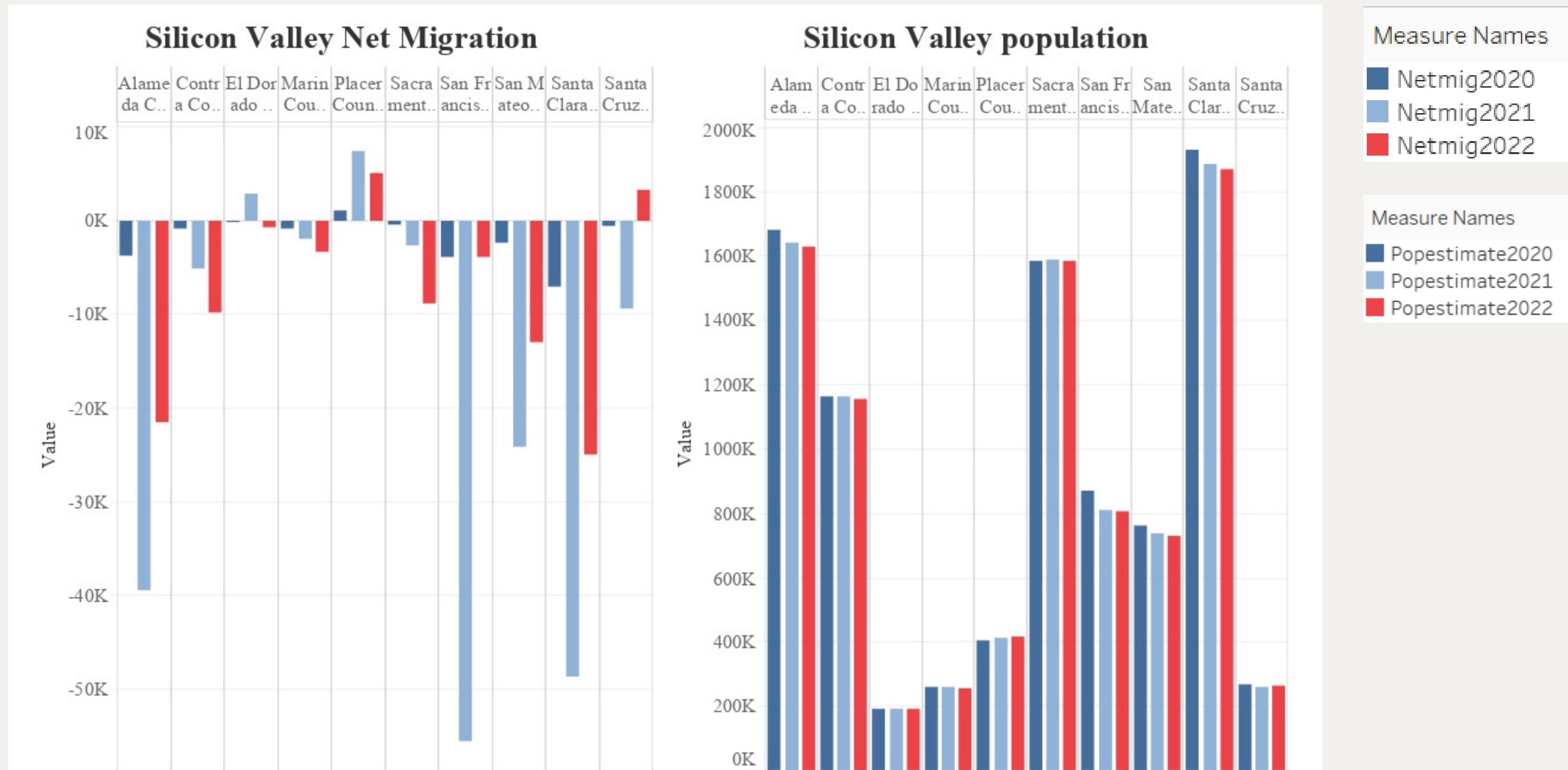


Counties Inflows Compared to Outflows and Income



Focus on Silicon Valley and South California

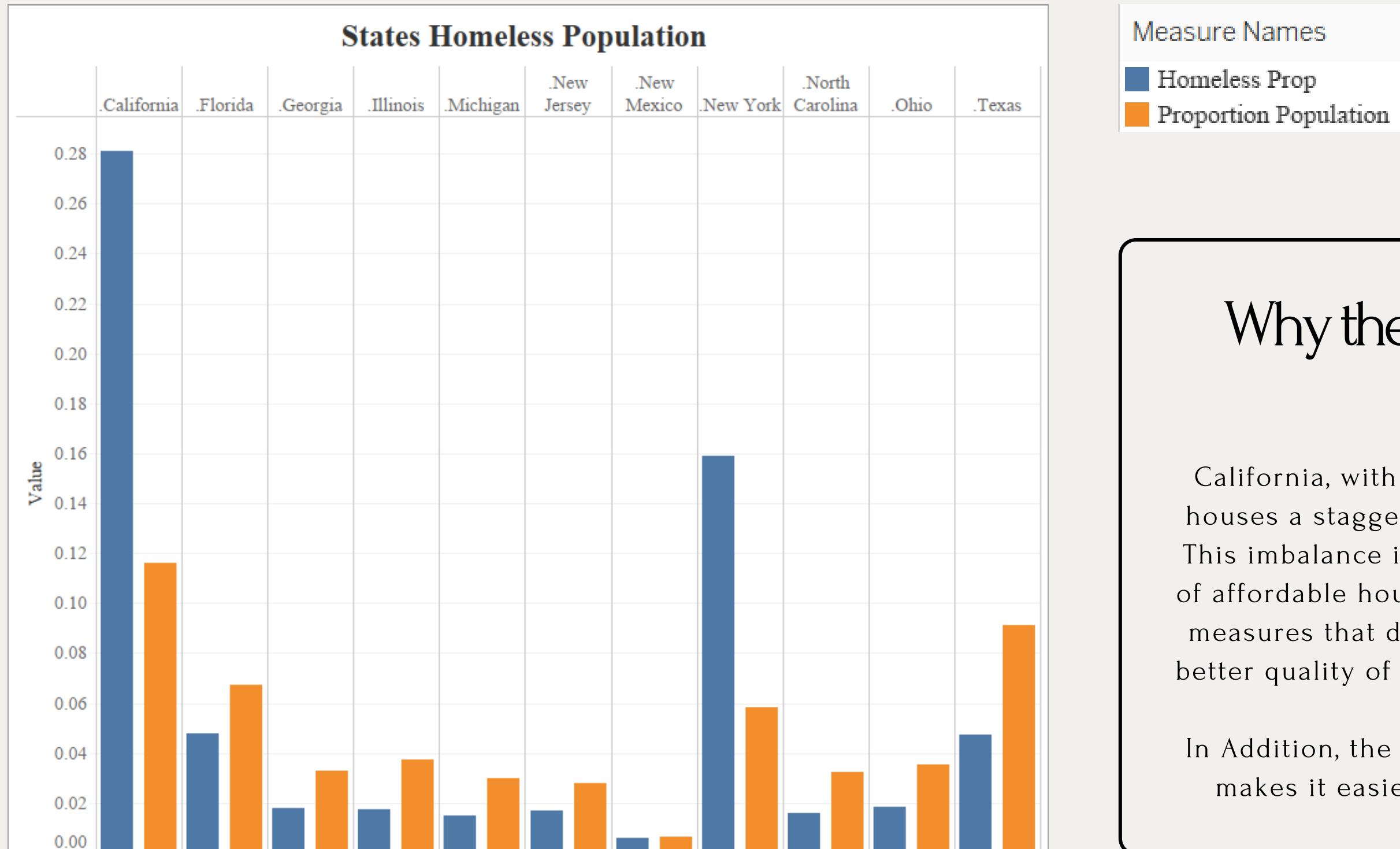
Silicon Valley Net migration and Population Data



Why the data is behaving as it is shown?

In 2021, many people left California for a range of reasons. The pandemic pushed people to move, and with remote work becoming the normal and common thing, living in states like Florida, Texas, or Nevada with no income tax was more financially beneficial for people. Plus, California's strict lockdowns, similar to those in Italy, led to lots of businesses going under. Rising crime rates also made people feel unsafe, so they packed up and left. Overall, these factors drove a noticeable exodus from California.

States Homeless Population Data

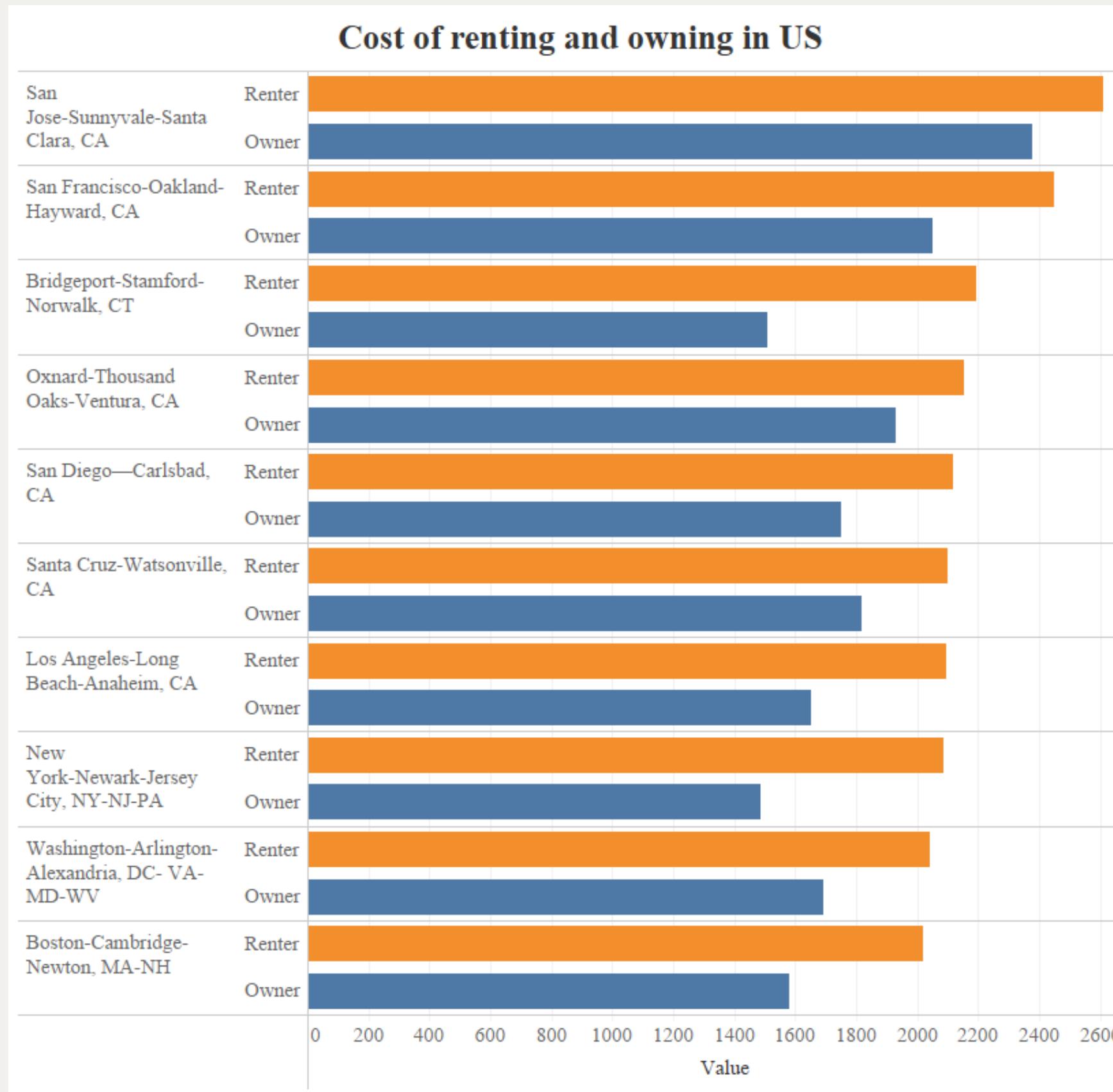


Why the data is behaving as it is shown?

California, with just 11% of the total US population, houses a staggering 28% of the country's homeless. This imbalance is largely due to factors like the lack of affordable housing, income inequality, and welfare measures that draw homeless individuals seeking a better quality of life. Similarly, New York experiences similar challenges.

In Addition, the climate in California is warm, and it makes it easier for homeless people to survive.

Cost of Renting and Owning in the US Data



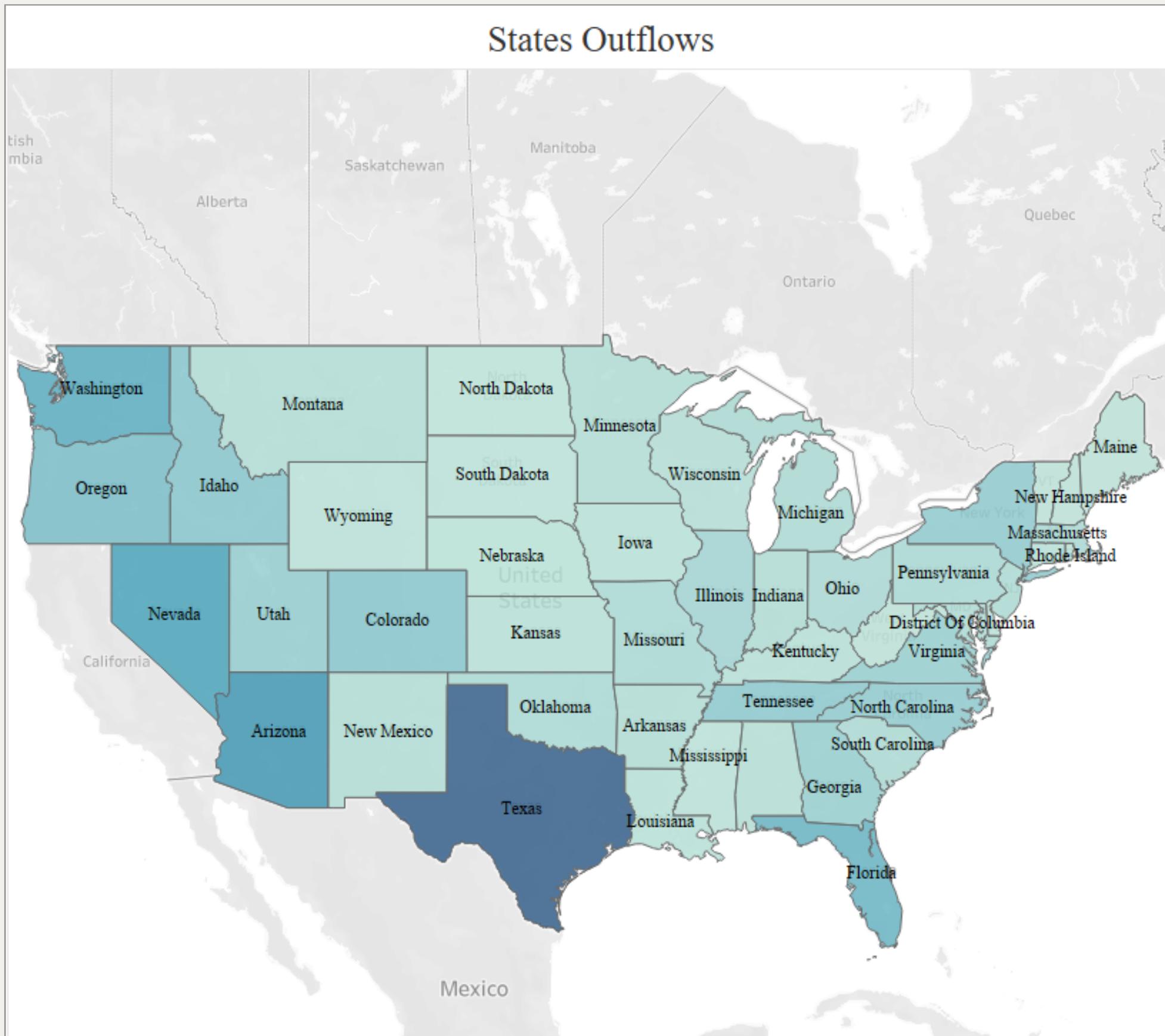
Measure Names

- Renter
- Owner

Why the data is behaving as it is shown?

Housing in California comes with a hefty price tag due to several reasons. Firstly, places like San Francisco are squeezed for space, making it tough to build new homes (one more example can be Manhattan). Then there was the law that actually favored people already living in California, making it harder to build new apartments. This law was called Proposition 13. Plus, because California's got a lot of money flowing around, housing prices naturally shoot up. All these factors make finding affordable housing in the state quite a challenge.

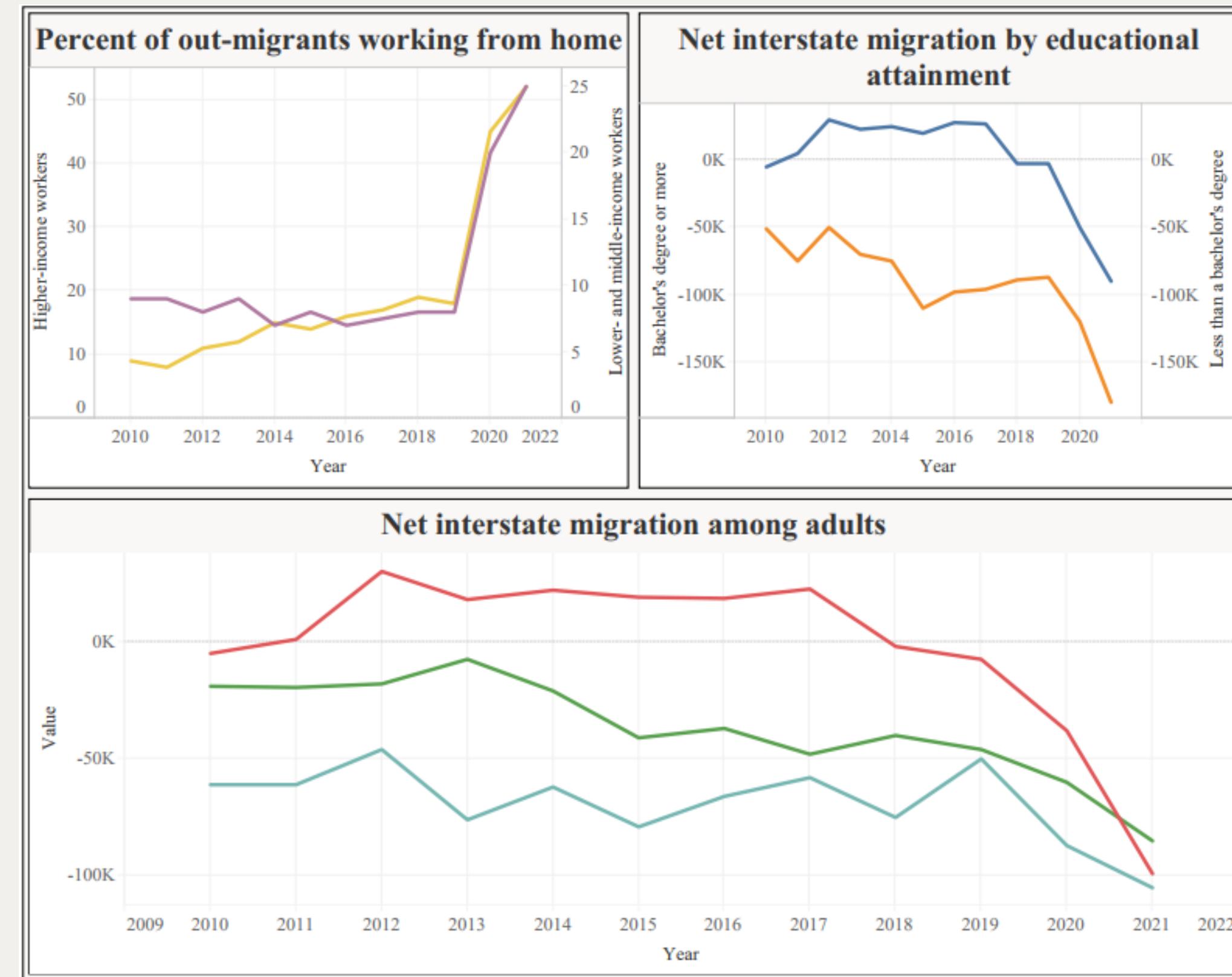
States Outflows Information



Why the data is behaving as it is shown?

Many Californians opt to move to Texas due to its appealing attributes. Texas's lack of state income tax, along with its more affordable housing market and relatively lax regulations across various sectors, makes it an enticing choice for those seeking better economic opportunities and a more accommodating environment for both personal and business endeavors.

Interstate Migration Data

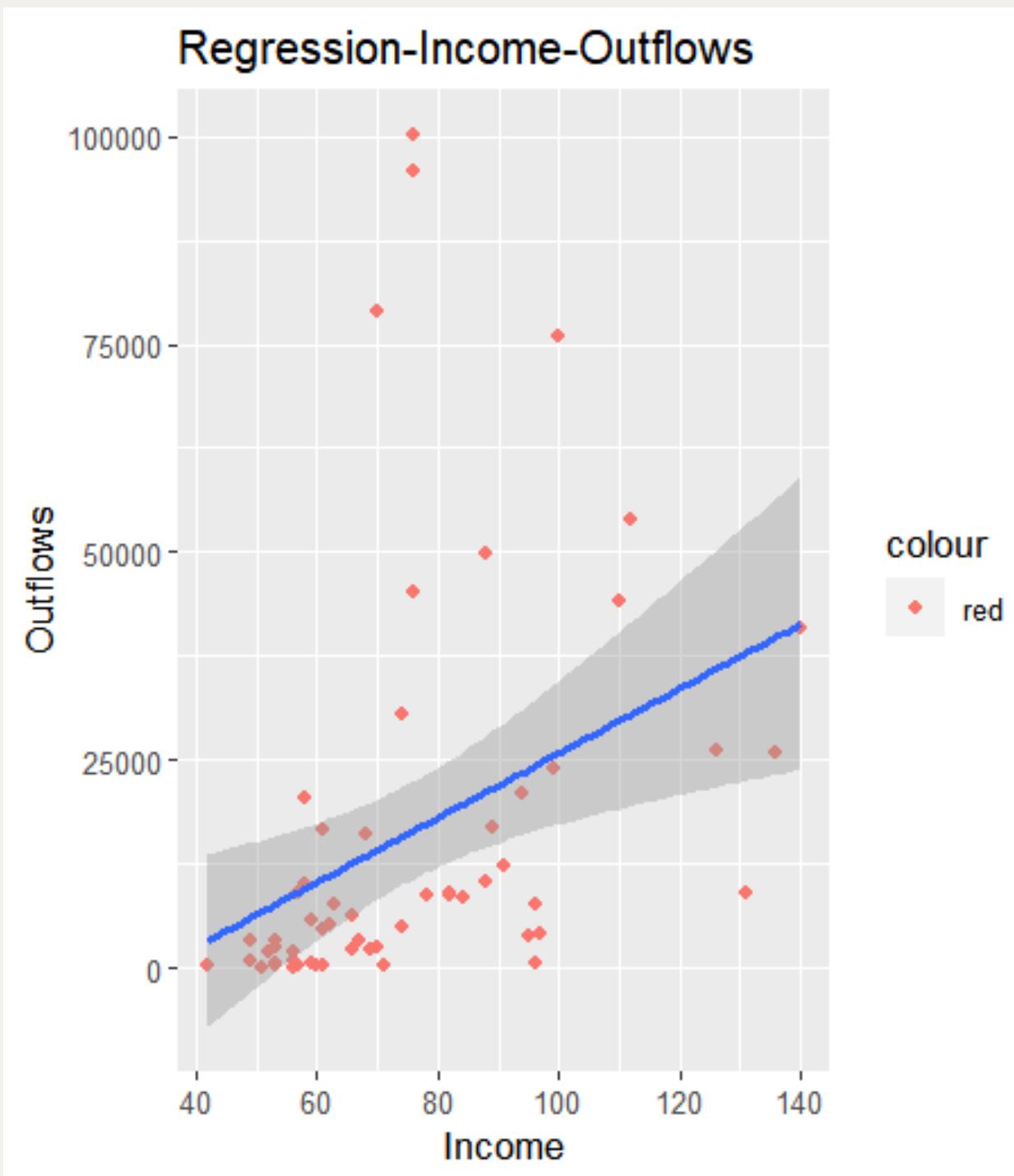


- Measure Names**
- Higher-income workers
 - Lower- and middle-income workers
- Measure Names**
- Bachelor's degree or more
 - Less than a bachelor's degree
- Measure Names**
- Sum of Higher income
 - Sum of Lower income
 - Sum of Middle income

Why the data is behaving as it is shown?

Across all levels of education, individuals left California in significant numbers, with a notable trend among those without a degree. The challenges exacerbated by the COVID-19 pandemic intensified this migration, although recent observations suggest a gradual reversal. Departure rates among individuals lacking formal education are particularly pronounced, driven by the increasing difficulty of paying all the bills without specialized skills or credentials. The advent of remote work, particularly advantageous for higher-income earners as previously discussed, facilitated the relocation of individuals from California.

Linear Regression



Insights:

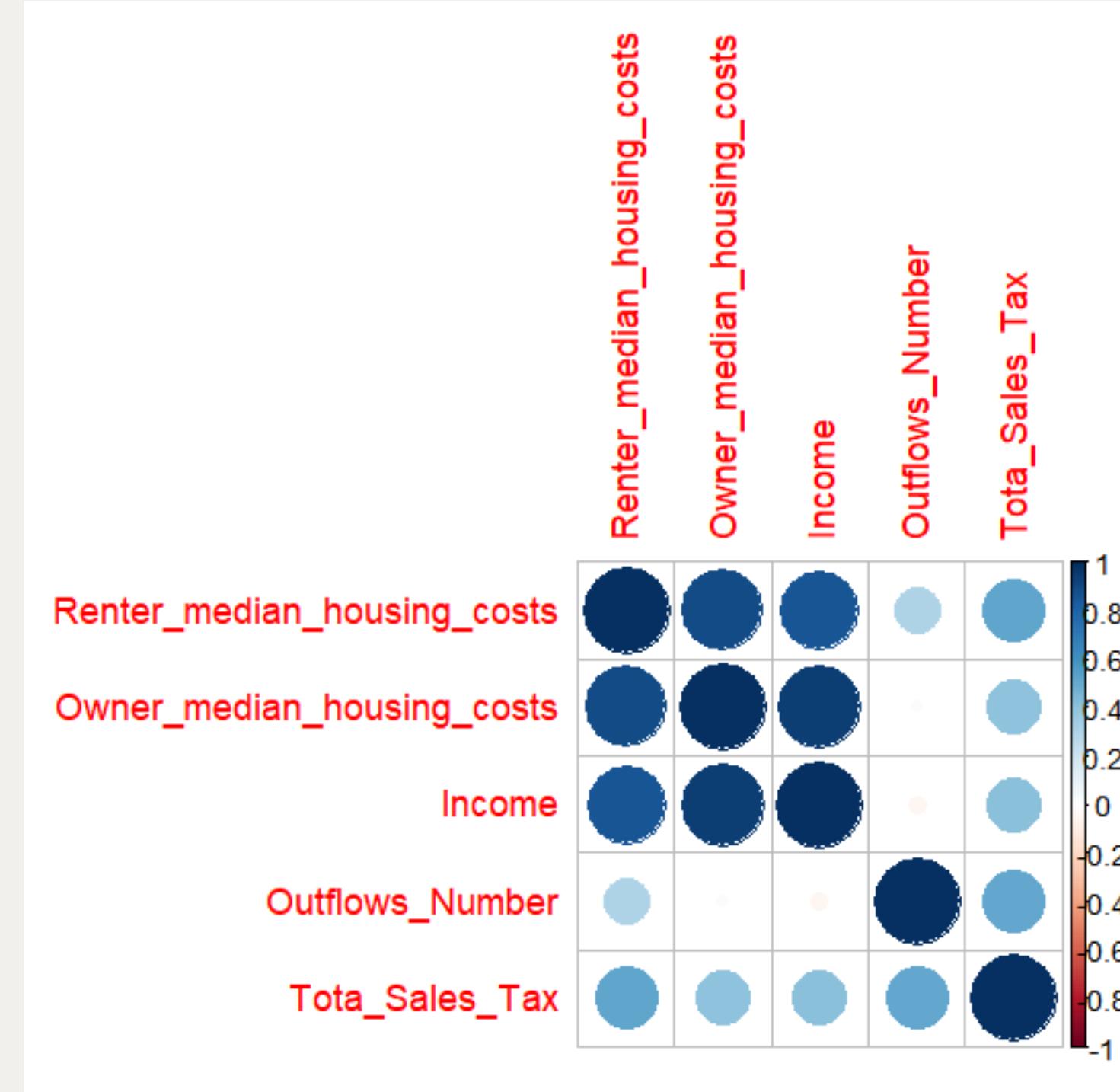
From the output we can see that we refuse the H₀ hypothesis of non-significance of the coefficient because the p-value is lower than the usual Significant values 0.05.
So we can say that, if we increase the income of 1 units the outflow will increases of 388.4 people and it is not possible, that if the income increases, more people will go away;
We will try to find some regressors, that can explain the cause of outflows.

Correlation

Insights:

In our investigation focusing on Silicon Valley counties, we delved into four key variables: renter median housing cost, owner median housing costs, income, and tax rate. Through correlation analysis, we discovered strong correlations among three of these factors. This indicates that including all of them in a single regression model may lead to multicollinearity issues and potentially skewed results.

Furthermore, our examination revealed an unexpected finding: there's a lack of significant correlation between outflows and both owner median housing cost and income. This goes against the common assumption that higher incomes and property values would naturally attract people rather than repel them. This observation suggests a need for deeper exploration into the dynamics influencing migration patterns within Silicon Valley counties.



Linear Regression

```
Call:  
lm(formula = Outflows_Number ~ Renter_median_housing_costs +  
    Owner_median_housing_costs + Income + Tota_Sales_Tax, data = data_regression2)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-34630 -7882   2041  12048  24566  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) -137829.61    78593.96  -1.754  0.1176  
Renter_median_housing_costs   108.93      55.25   1.972  0.0841 .  
Owner_median_housing_costs    -45.74      85.66  -0.534  0.6079  
Income          -1027.23     928.35  -1.107  0.3007  
Tota_Sales_Tax     1478038.29  983231.91   1.503  0.1712  
---  
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 20830 on 8 degrees of freedom  
Multiple R-squared:  0.5764, Adjusted R-squared:  0.3646  
F-statistic: 2.721 on 4 and 8 DF, p-value: 0.1064
```

Insights:

As we said in the previous slide, we took 4 new regressors. Second attempt to make a multi-linear regression resulted in low Global Significance, which indicates, that the model is not significant for the dependent variable; None of the regressors is significant for the analysis, the renter housing cost is significant at 10%. In this case we can notice that the income and the owning housing cost are negative, which makes sense and make the model realistic.

Recommendations and insights



Based on conducted data analysis, we have come to the conclusion, that even though the regression analysis showed, that the indicators are not that significant due to lack of the available data for each County, many of the aspects, that we have analyzed, can explain what is happening in California. For instance, Los Angeles and San Francisco are the top cities for the organized retail crime, and California has the widest population of homeless people.

Gentrification: instead of happening in just a small city, it is happening in the whole California, so as we have seen previously, lower and middle income people are more likely to leave California, rather than people with higher income. Based on our findings, we would suggest to soften the regulation of housing and “force-increase” the building of houses in California. Dealing with retail criminality might be a solution, as well as reglamentation of the gentrification.

