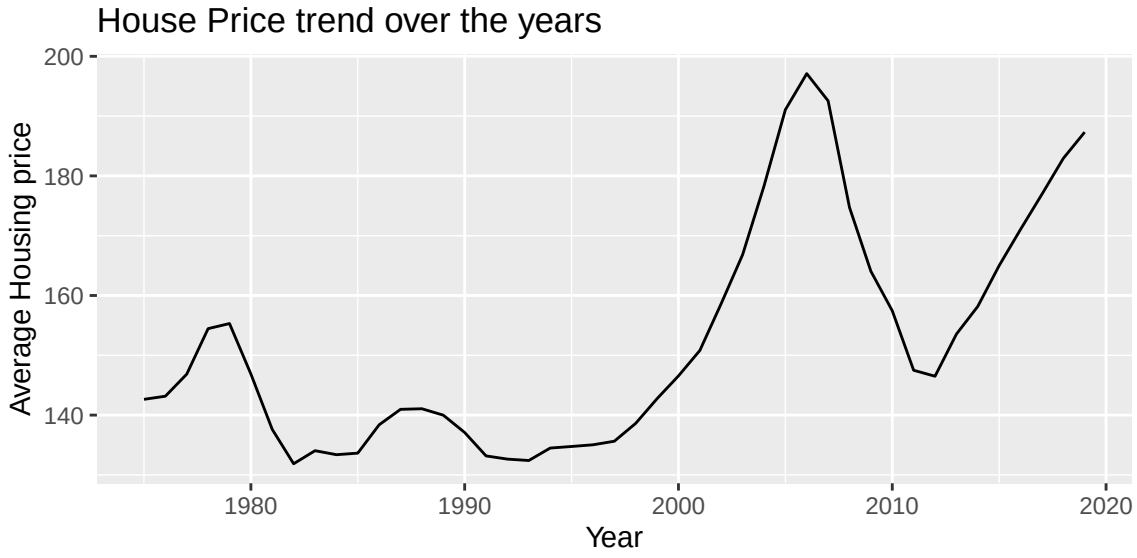
Mini Project 1

Team Nevada

Loading Data

Question 1:House Price Over Time

The original Data from freddie Mac had house prices of all the states for ecah moth of the year. We have taken an average for all the States in a year and then plotted the graph to observe the trend for all the 51 states. For observing the overall trend in house prices from 1975 to 2019, we have taken an average of all the prices in a particular year.For trend in house Prices with in a region we have taken the average of all the prices for a Region and Year. We have adjusted the house prices for inflation as per the December 2019 dollar value for all the plots.

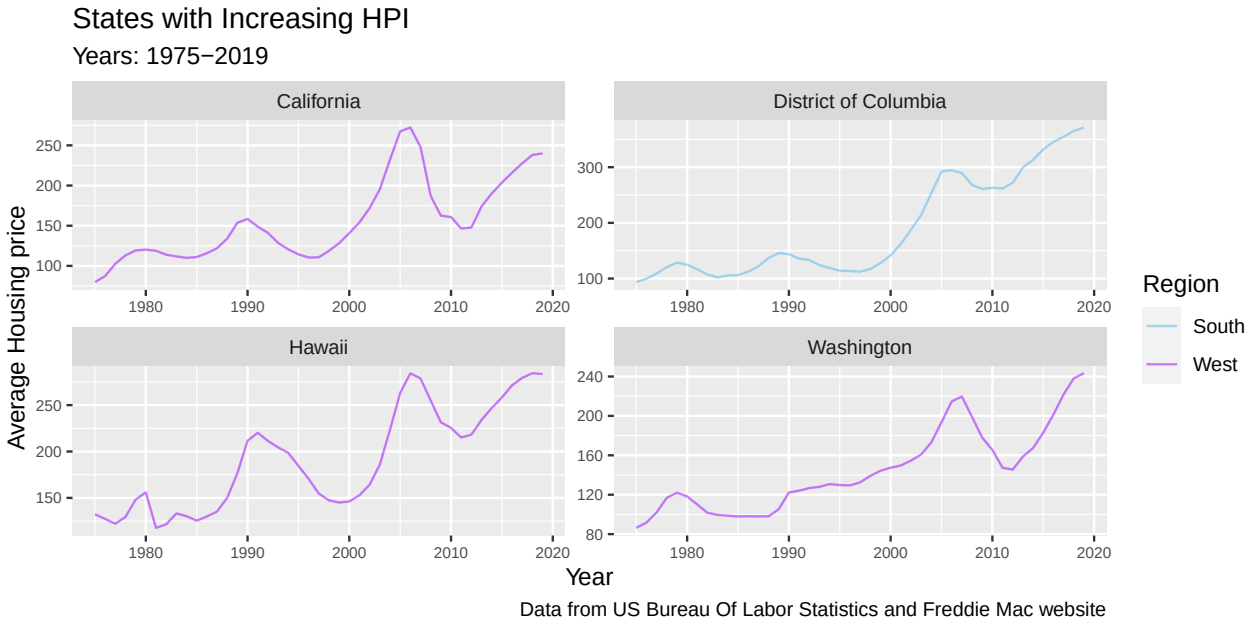


Overall Trend:

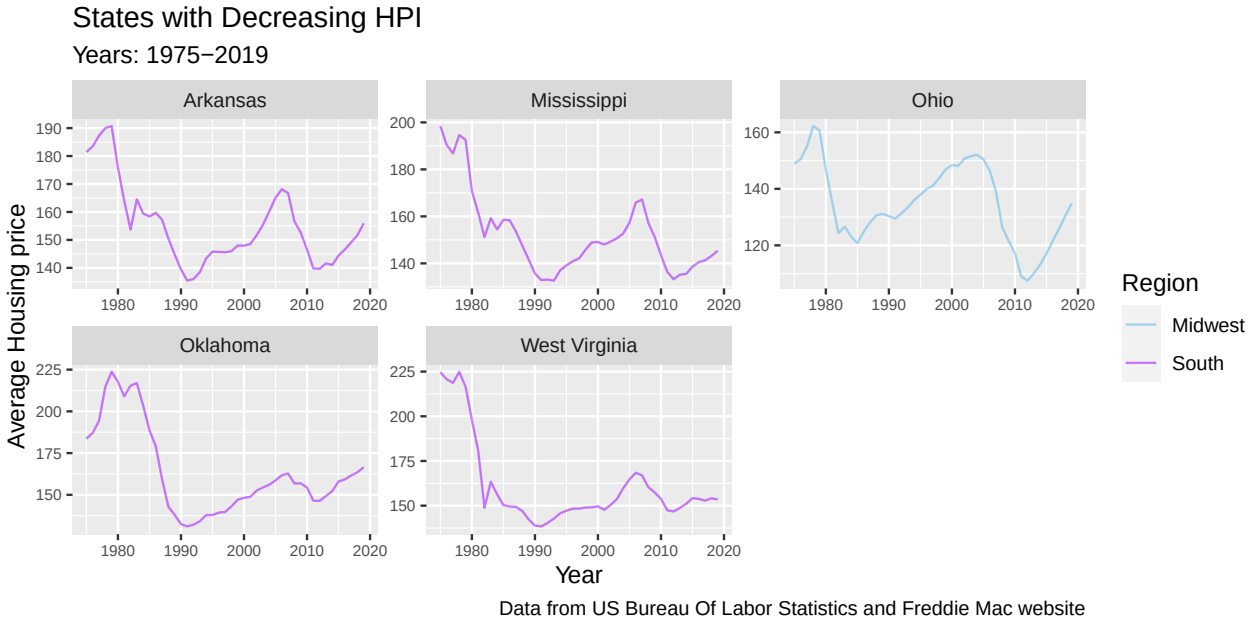
We can see that the Plots shows a dip during the early 1980's and 1990's recession as well as a decline during the great depression of 2008.

We can see that the states in the Same region follow the same pattern with a few outliers. We can easily observe that Hawaii, District of Columbia, Washington and California are states that have increasing house prices. Similarly, states like West Virginia, Mississippi, Arkansas, Ohio and Oklahoma have decreasing prices.

States with highest increase and decrease :

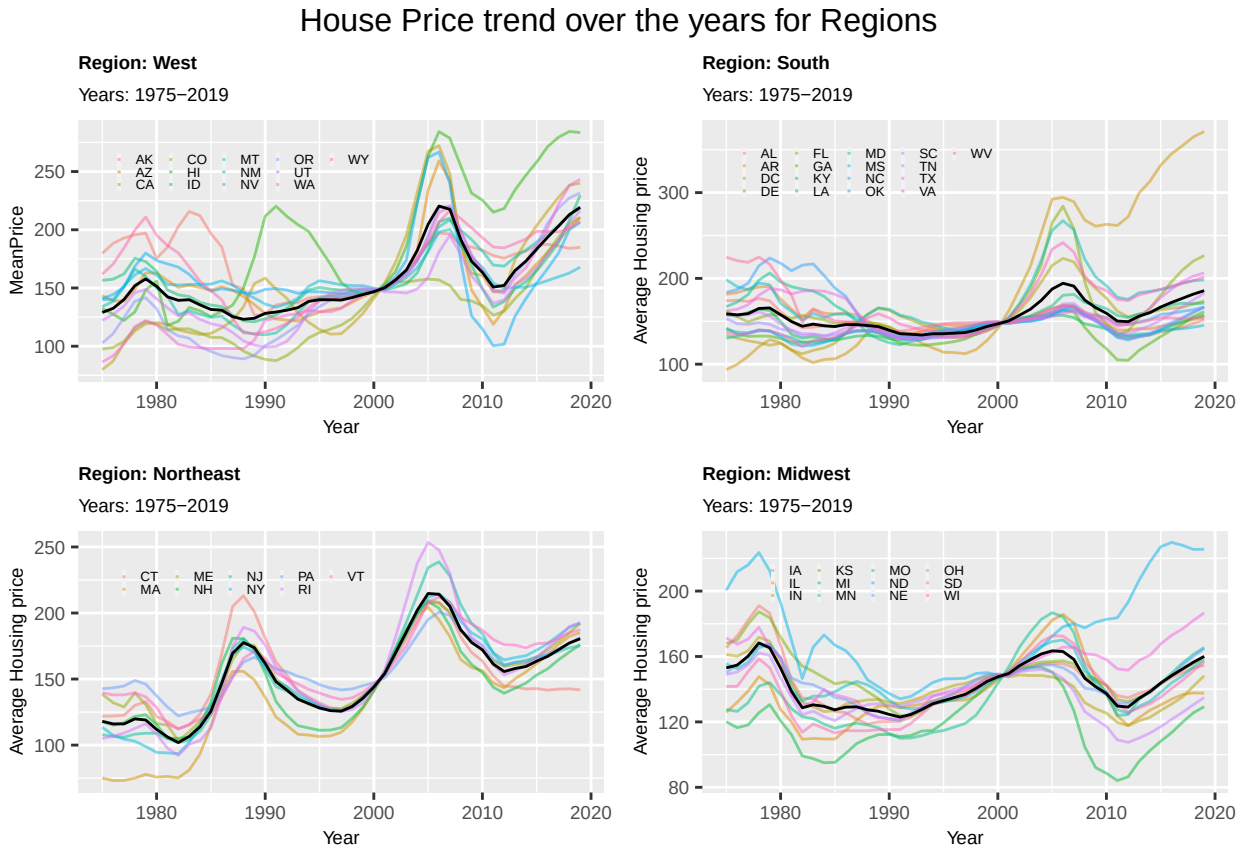


We see that District of Columbia have a constant increase in prices after 1995. Washington had steady increase in prices from 1989-2004. Hawwaii had 3 major dips in the time peroid of 1980-1984, early 2000's and 2010-2013. For California we see an sudden increase between 2000-2006, which could be due to the increase in IT growth. All these states show an increase after the depression of 2008



States from the South region follow the same trend of a decline during 1980-1990 and then a increase in prices till 2006. For example, Oklahoma and West Virginia shows a constant decline in prices except for a slight increase between 2000-2004.Ohio shows that it reaches a peak during early 2000s, but since has reached a low during early 2010's. We can observe that the District of Columbia has seen the highest increase in prices and State of West Virginia has seen the highest decrease in prices.

Region-wise Analysis:



Do the typical patterns vary between the four regions (Northeast, Midwest, South, and West)?

The Black line is the average of a Region. We can see that all the Region have a different rate of increase in house prices. Midwest and South Region have a plateau during the 1980-1990. Midwest has a sudden dip during 1990-1992, after which it has almost linearly increased till mid 2000s. The south Region has a slower increase after the 1990 dip. The West Region has a consistent increase in prices after the 1989 dip. The Northeast region shows an overall upward trend. All the Regions show an increase during 2006-2010 period after which there is a constant decrease due to the economic depression of 2008. Similarly all the regions show a decline during the period of 1978-1983.

Have changes in prices within each state mostly followed the same basic pattern, and are there outliers to that pattern?

We can see that all the states in the same region follow a similar trend except for a few outliers.

West: There is an overall increase for the time line of 1975-2019. The common trend that the states in this region follow is the shift in prices after 2000. All the states a dip during the 2005-2013 and an increase before it from 2000-2006. States like Hawwaii, California and Washington have a steep increase after 2014

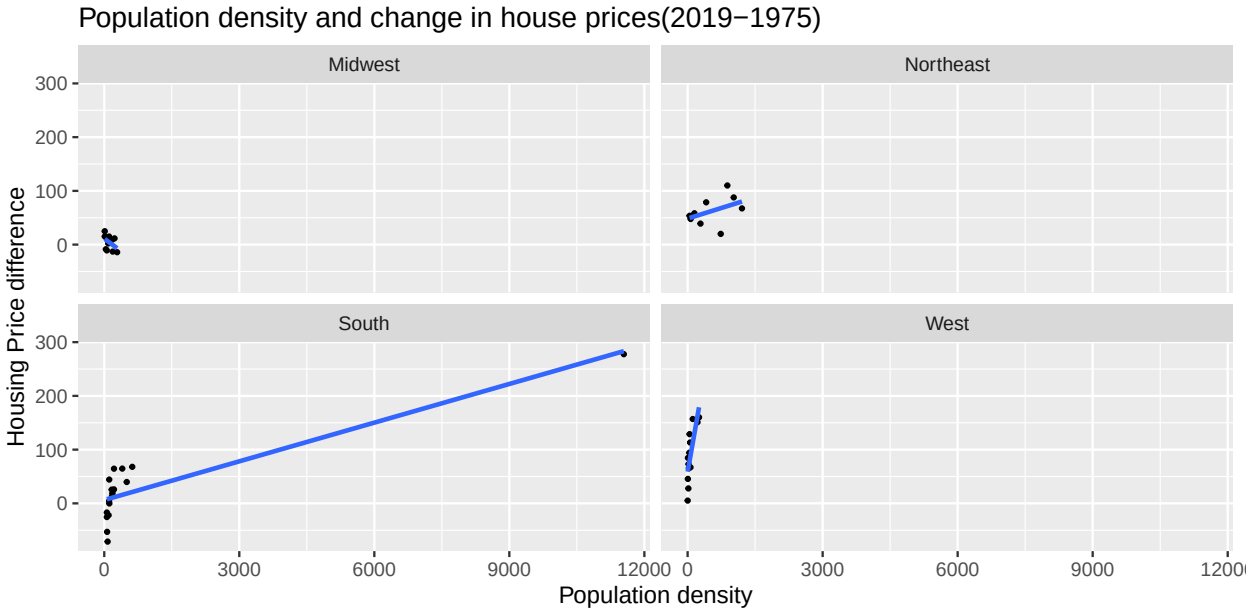
South: Apart from District of Columbia,all the states in this region have followed a similar price increase after 1989. District of Columbia is the highest increase in prices compared to all the US States, going over 300.

North East: All the states show a similar two peak trend during 1985-1990 and 2004-2006. House prices of Connecticut show a decrease which is not the pattern followed by other states in this region. Rhode Island reaches the highest peak in this region during the 2005-2010 period.

Midwest: All the states follow the similar pattern except North Dakota. Prices have increased for all after 2010. All states have a peak near 1987-1989 and another peak at 2004-2006. Minnesota shows a sudden decline from 2004 to 2011 which is not in tandem with all the other states All the states in a region show a similar changes during the period of late 1990' and early 2000's. But the pattern differs when we are comparing states of different regions.

Question 2: Population density and changes in house prices For explaining the change in house pricing since 1975 from the present-day population density, we are finding out the difference in housing prices from 1975 to 2019 for all the regions by taking annual average of the housing prices of both years and then finding out the difference by using price value 2019-price value 1975.

Plot to see the change in population against population density

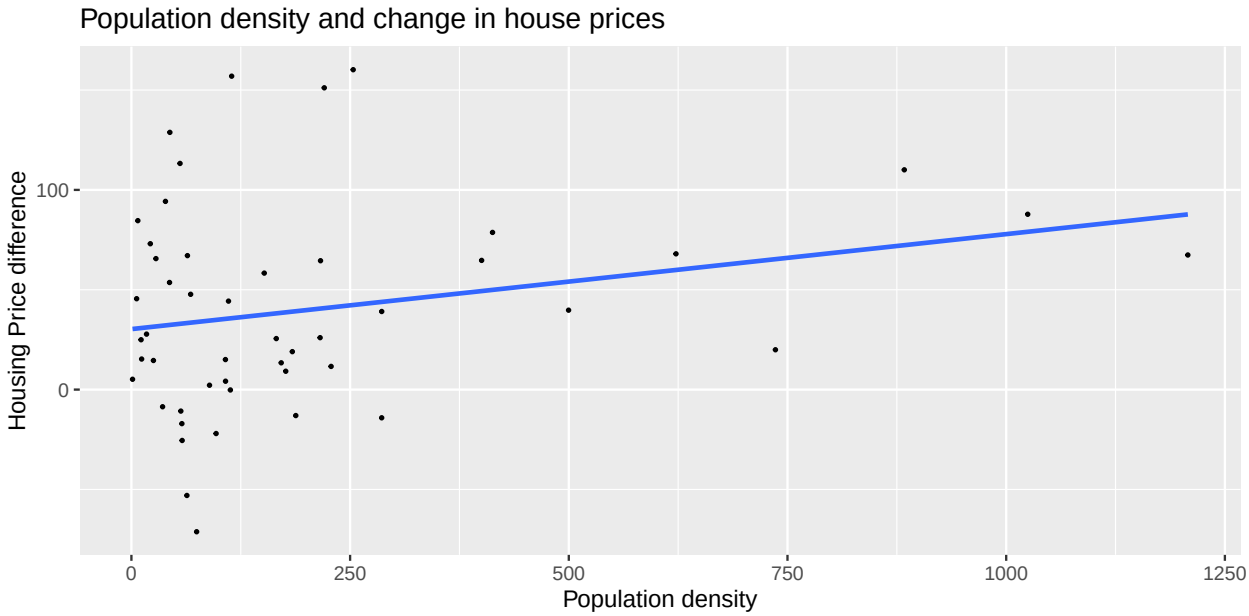


In the above plot of south region area,

The plot for Population density and changes in house prices shows different trends for all regions. We can compare our X-axis vs Y-axis and see how price difference for 1975 and 2019 is showing different trends. we can see a gross outlier(District of Columbia) which makes us difficult to analyse rest data. The population density for DC is more than 11K. If we want to study south region, we dont see any pattern for the rest of the data because of this outlier. Blue line shows the linear regression line, which finds the trend in data by averaging the neighborhood values. In the South region, if we remove this outlier then we can see a meaningful pattern for the rest of area in below plot. and we can see the same density in our graph too.

This graph again shows different trends for different regions. In this graph, we have linear model to see the trend between housing price and Population density. Few of the general findings for this graph: 1. South, west, Northeast data is showing increasing slope, however Midwest data is showing decreasing trend for Housing Price with increase in Population Density. 2. If we see the slope gradient for all the four regions then we can say West slope gradient is having highest gradient, however, Northeast is showing a little parallel line to Population density which means lowest gradient. 3. Midwest: In this region, plot is showing a decreasing trend and if we see our data too, many of the states price difference is showing negative values as their density is increasing. Ohio has highest density; however, diff is in negative value. There are little bit fluctuations in data which is not readable in data because second highest density data is Illinois but the price diff is positive. However, our plot is mostly showing negative slope. There could be some other factors too, but still, we can say our Midwest data is having negative association between these two variables. 4. NorthEast: Data and plot both are showing same trends. In data also we don't see much difference in association with these two variables. That's why we have almost parallel straight fitted line wrt X-axis. 5. South: In this region, Data shows density with higher population shows higher difference in housing price. We can see that in our plot too that highly dense population has higher housing prices too. 6. West: This region is showing drastic change between these two variables. A sudden increase in density shows steep change in price difference.

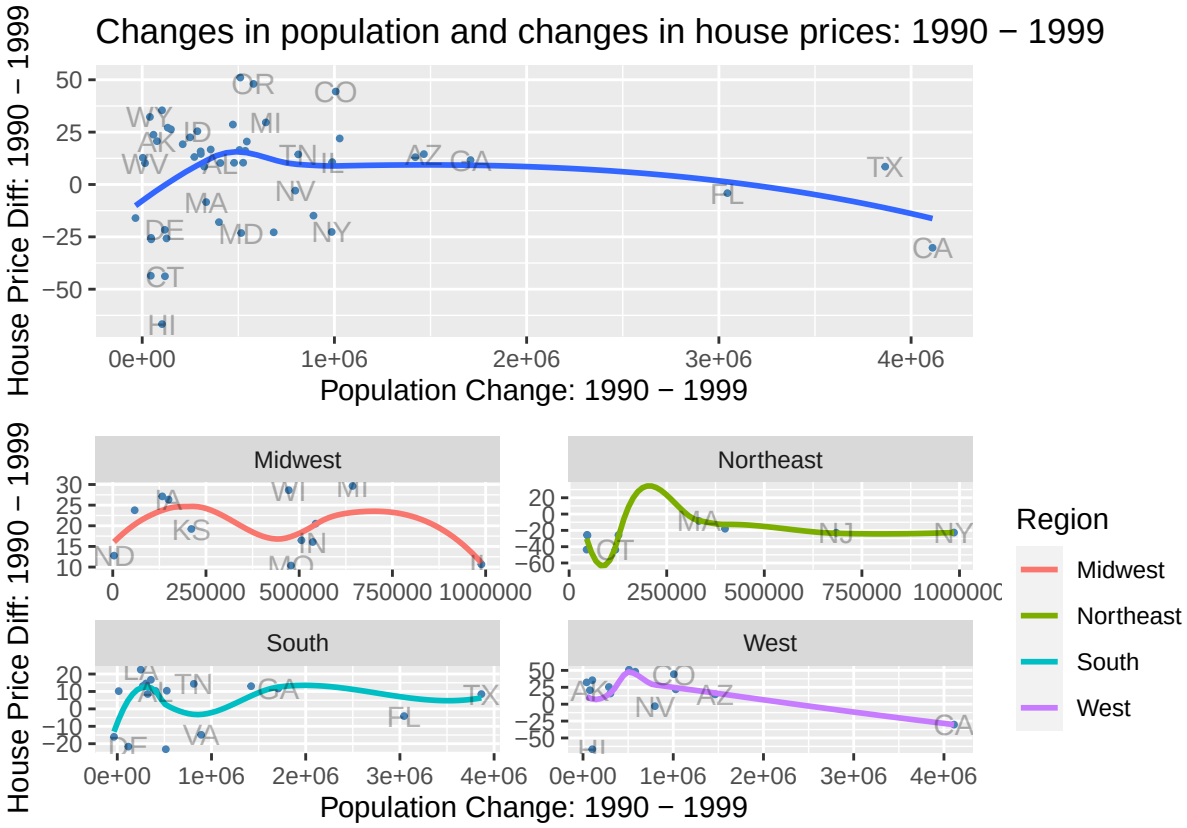
Below plot gives us a overall scatter plot for all the regions including all states. As we can see few data points like areas of west region and DC are having higher population density. If we see the overall picture we can understand that these two terms (population change and housing prices) are correlated. We can check this section in detailed in Q3 for each decades.



Question 3: House price variation with change in population by decade

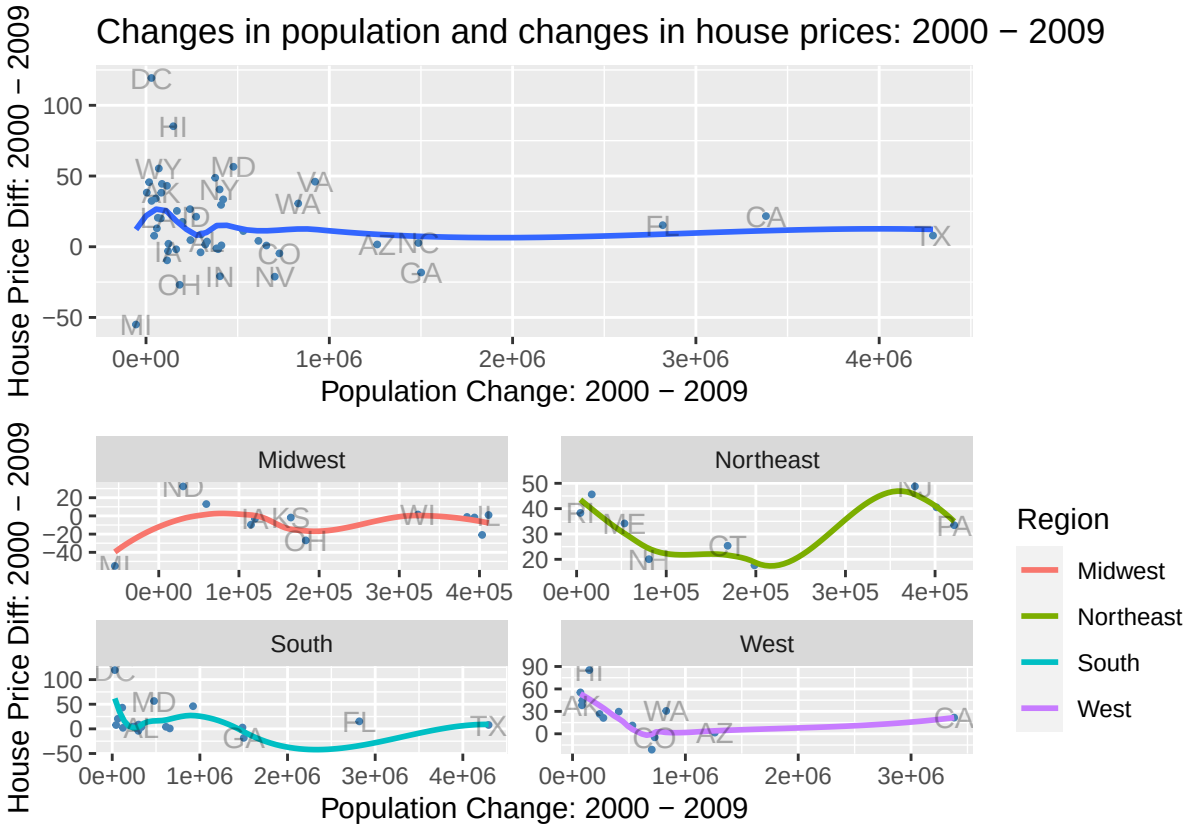
Initially, we filtered the population data to remove all the columns except state and population of 1990,2000 and 2010. Using Census API key, we obtained the population of latest year. To understand the change in population from decade to decade, we have subtracted the population for each decade and added it as new column. We adjusted the house price of all years to account for inflation by multiplying with CPI index. For each year, we calculated the mean house price to plot graphs.

1990 -1999:



For the decade 1990-1999, If we observe overall variation, population was decreased for DC and so is price. For the states such as HI, RI, OT we can see that even with the slightest increase in population house prices have decreased by a huge scale especially for HI. For the state FL and TX, we can see tremendous increase in population, but house price remained almost same. For OT even with little change in population, house prices have increased drastically.

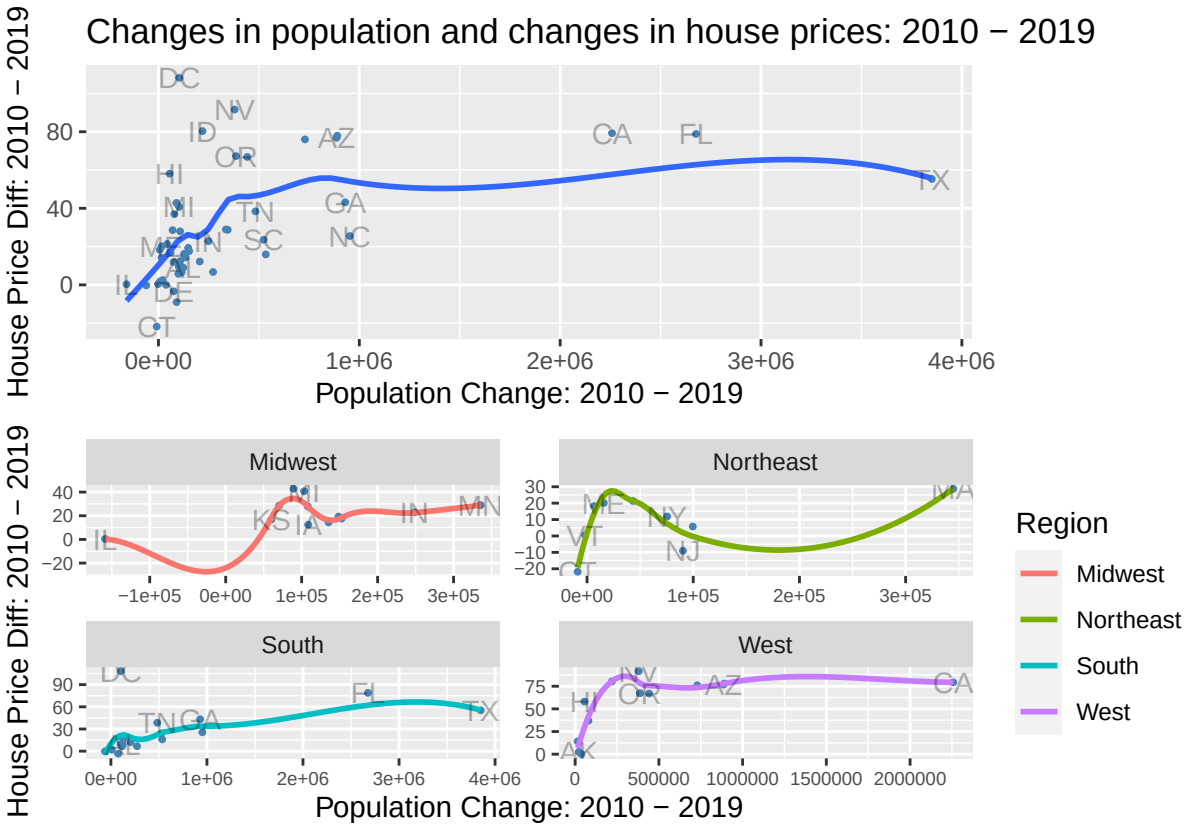
Again, if we individually examine various regions, particularly *Midwest* doesn't follow the same pattern as Overall trend. For this we could see there is no decrease in price with the change in population, except for ND and MO we can see that there is increase in house price with increase in the population. Even for ND and MO we do not see any decrease in the price, For ND there is slight increase in price with the constant population whereas For MO price is constant although there is a considerable increase in population. For all other three regions, we can see the same trend as that of overall variation.



2000 - 2009

For the decade 2000-2009 for DC unlike the first decade, there is huge increase in the house price with almost no variation in the population. For the state HI, unlike previous decade with the slight increase in population, we can observe massive increase in house price. For the states FL and TX, trend remained same as last decade. In this decade, it can be observed that for MI, even with the slight decrease in population, there is a considerable dip in the house price.

Again, if we individually examine various regions, particularly *Northeast* region, there is no negative house price difference unlike the overall trend. For the *Midwest* region, no other states except ND have seen a tremendous increase in population. In the *South*, the sates LA and AL there is almost no variation in house price although there is slight increase in population.



2010 -2019

For the decade 2010-2019, the state DC follows the same pattern as second decade with the huge increase in the house price with almost no variation in the population and HI also follows the same pattern as we can see with the slight increase in population, we can observe massive increase in house price. The states FL and TX followed the major deviation from previous decades witnessing an increase in house prices with increase in population.

Again, if we individually examine various regions, For the state MI in *Midwest* region we can say that both population and price have increased tremendously compared to previous decades. *South* Region has seen the highest increase in the house price amongst all other regions although there is very less increase in population for most of the states. In the *West*, particularly the states NV, AZ, and CO there is huge increase in price with slight increase in population unlike the first decade.

Conclusion: Going by Region-wise analysis, For **West** Region, the population of this region has not been increasing, however the house prices have been high for almost all 3 decades we are observing. **South** Region shows stable changes in population and hose prices for all the 3 decacdes. **NorthEast** region shows drastic changes for both house prices and population. For the decade of 1990-2000 the house prices didn't go up even if there was an increase in population. After 2010, the house prices have seen drastic increase. **Midwest** show a increase in population for all the decades, but the house price doesn't increase monotonically. There is an increase in house price for the time period of 1990-2000 and after 2010, but a decrease in prices for decade of 2000-2010, which could be due to economic slowdown. So, it's safe to conclude that there is no fixed relationship between House price and population change as it varies differently for different states throughout the three decades. It can be inferred that variation in house price does not only depend on the population change but there probably are some external factors that influence it.