AAKASH AHUJA | AYUSH BHATIA | RUSHABH SHAH

PROBLEM STATEMENT :

How have house prices in U.S. states changed over the last few decades, and are changes in prices related to population in some way?

INTRODUCTION :

We explore the fluctuation in house prices in the US over the past 4 decades using the **Freddie Mac House Price Index** data and determine the factors affecting the changes in the House Price. We begin by inspecting house prices against the tide of **time** and like most things time has had its fair share of impact on the house prices. Then we move on and evaluate the changes in house prices against various other factors namely: **Geography**(Regions and States), **Population density and Population change**.

HOUSE PRICES OVER TIME

We begin by identifying how the house prices have changed since 1975.
So our first task is to adjust this data for inflation!
We have month-wise house price indices and the consumer price index of the US which can be used for adjusting inflation. Adjustment for inflation is done by using the following formula:

$$New\ Price\ = \frac{Original\ Price}{CPI}\ \times CPI(year(2000))$$

The inflation-adjusted return is the measure of return that takes into account the time period's inflation rate.
We examine the relationship between house price index with time (1975 -2019). (Figure 1)

Figure 1 : HPI vs YEAR
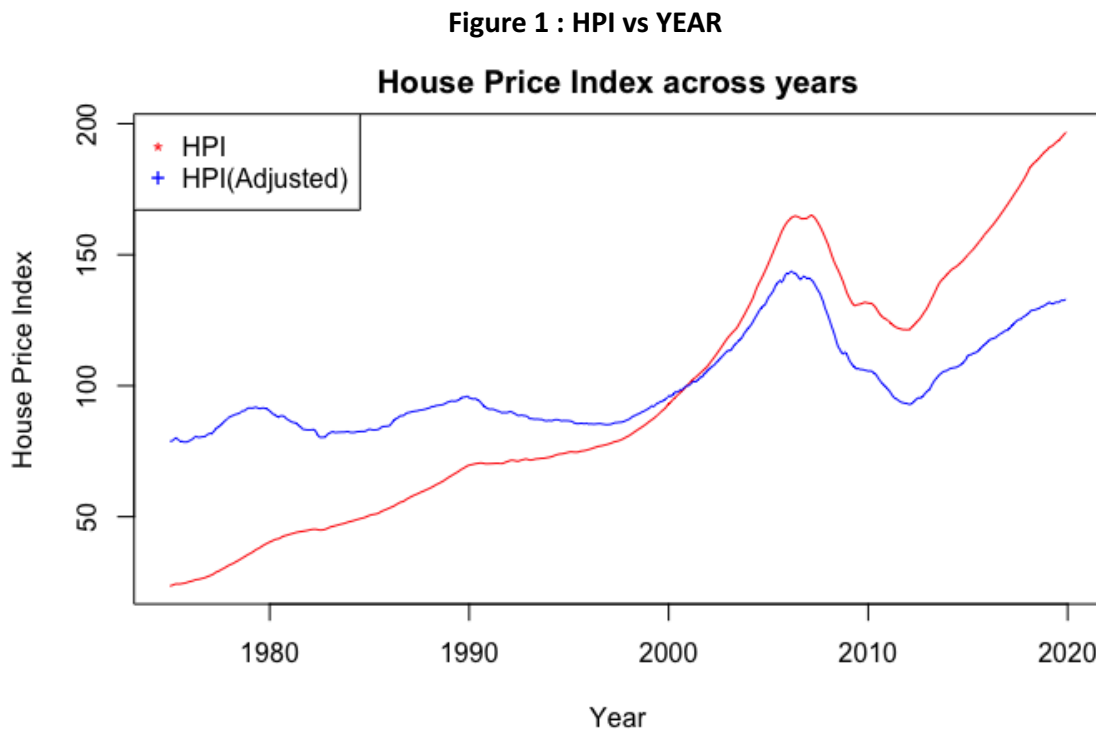
We convert the HPI data into time series data and fitted an additive model on it.
After adjusting for inflation, it can be seen that the overall house prices increased over the span of 45 years(1975-2019). However, there is a sudden drop in house prices observed during the time period 2008-2010(Due to the great recession). Then the house prices start increasing again after the year 2012. The overall trend in house prices is increasing as expected.

CHANGES IN PRICES WITHIN EACH STATE

Now, we observe the variation of HPI in each state and compare them to identify whether they follow the same pattern(figure 2). Again, we can expect that the overall trend will be that of an increasing nature.

We observe from our data across these 4 decades that the average house prices have increased for all the 51 states. While the biggest increase has occurred in District of Columbia(6.11 %), the smallest increase in HPI was in West Virginia(2.51%) (Calculation in R Code in dataframe Mean_prices)

Figure 2: House price across states
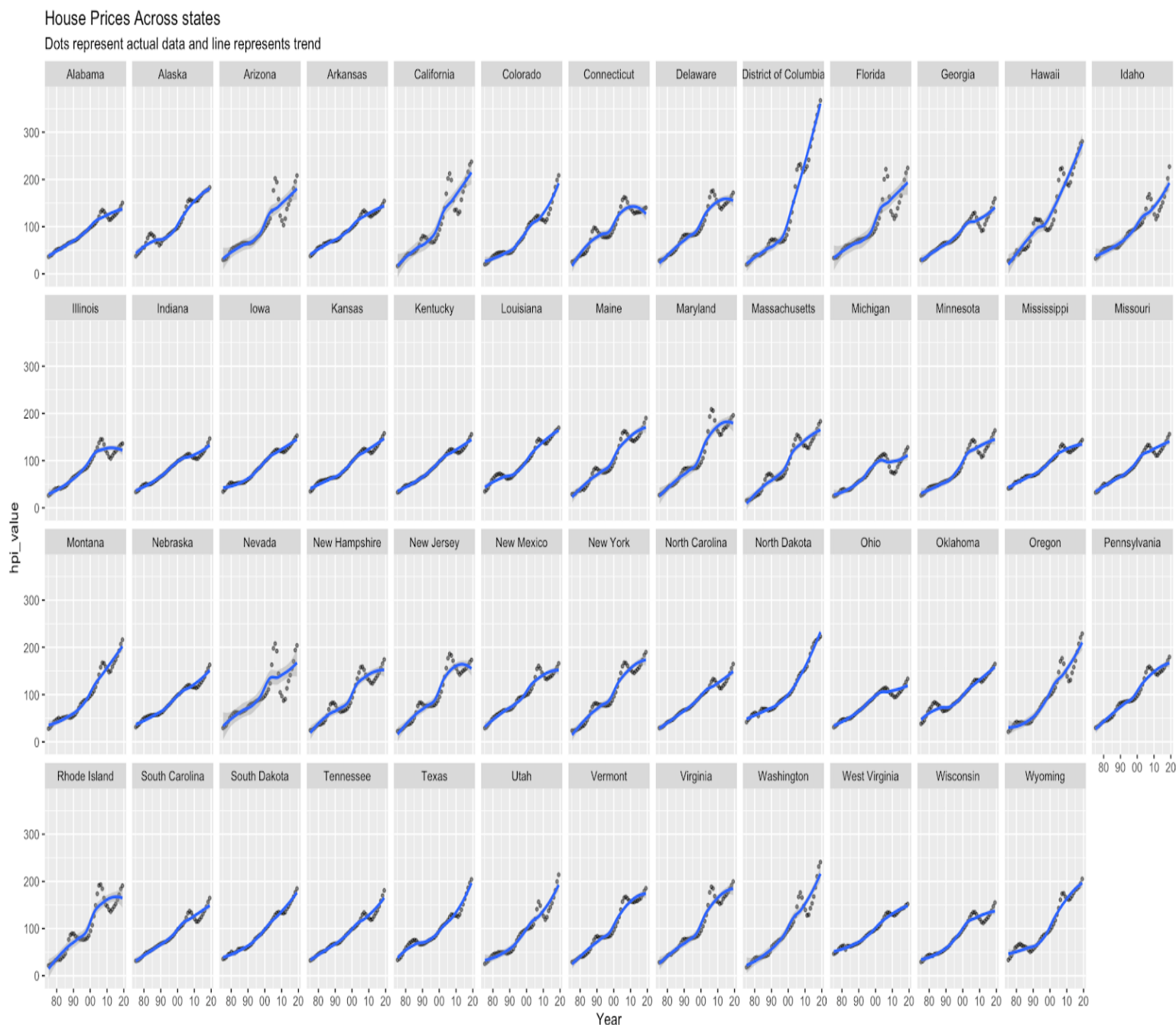
To observe the pattern between states we plot the data and fit a loess model on it to see the trend. It can be seen that almost all states follow the same basic pattern which is increasing HPI w.r.t year. The dots represent the actual data and the line shows the underlying trend.

There are a few states which do not follow a similar pattern as others and can be considered as outliers.

The biggest change of 24.65% in house prices was observed for Rhode Island for the year 1987 and the smallest change of -38% was observed for Nevada for the year 2009(Great Depression). Nevada and District of Columbia deviate from the basic pattern as the house prices keep on increasing even after the recession period of 2008-2010. These two states can be considered as outliers.
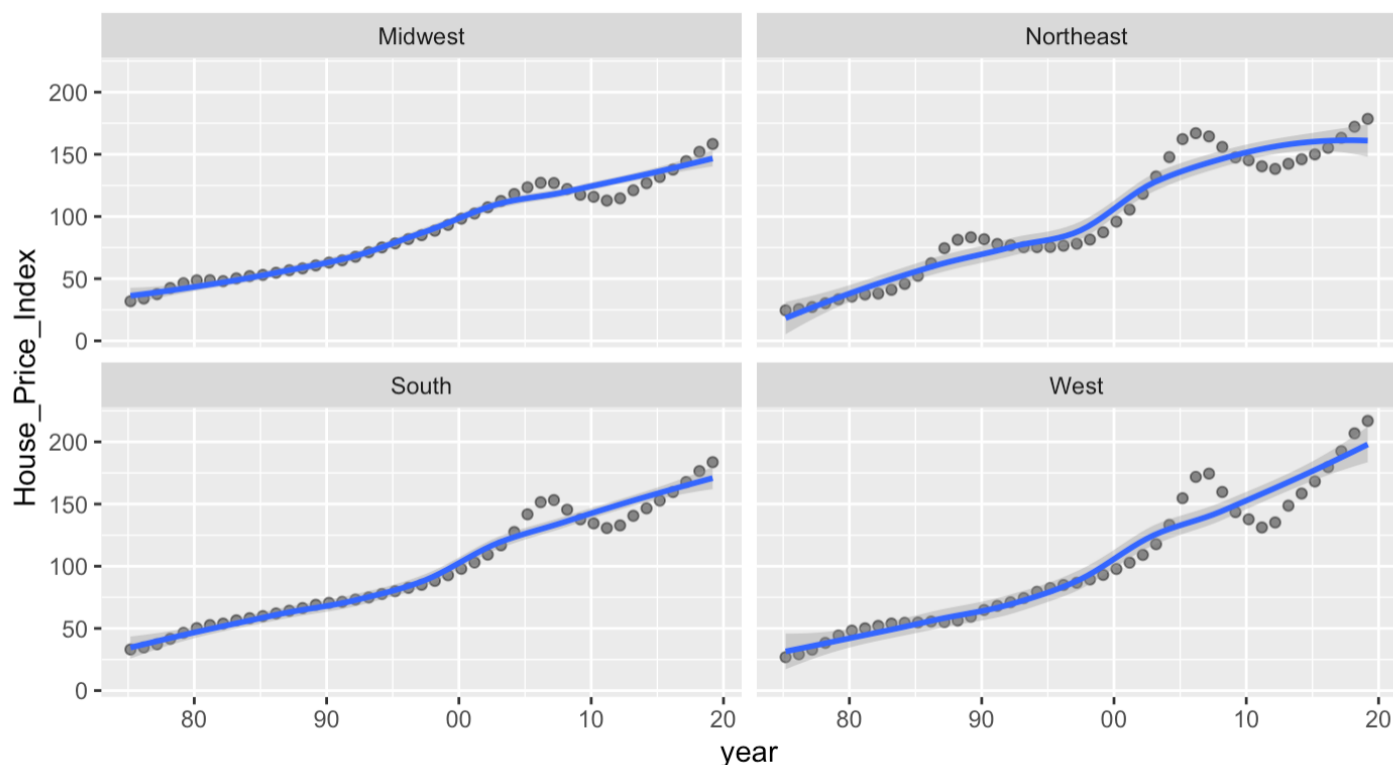
We observe state house prices data on a regional level by taking mean house prices of all states in a region.

Figure 3

House Prices Across US Regions
Dots represent actual data and line represents trend

The overall trend of house prices of different regions over the years is increasing.

The West region has the highest increase in house prices from 1975 to 2019 whereas the Midwest has the least. Northeast and West follow a similar trend while the Midwest and South follow a similar trend to each other in terms of changes in house price index across the years.

The general trend is that house prices increase over the period, with a sudden drop in house prices during the period 2008-2010. (Figure 3)
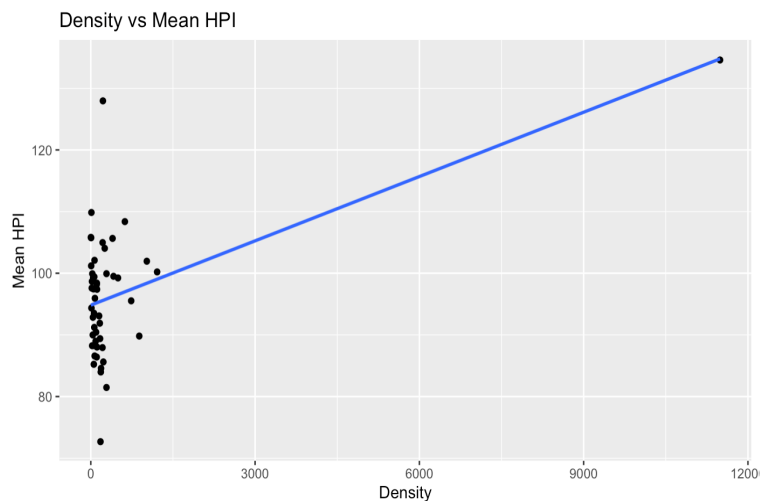
Here we fit a loess method.This is a nonparametric method because the linearity assumptions of conventional regression methods have been relaxed.The fitted points and their standard errors represent are estimated with respect to the whole curve rather than a particular estimate. So, the overall uncertainty is measured as how well the estimated curve fits the population curve.

**POPULATION DENSITY AND CHANGES IN HOUSE PRICES**

Population increases every year and has been increasing since 1975. The increase in population could have an effect on house prices. To identify whether population has an effect on the house prices, we compare the current population density with changes in house prices.

We take the current population density of each state and compare it with the mean HPI of states from 1975 - 2018.

Figure 4   i) Density vs HPI with outliers   **ii) Density vs HPI without outliers**
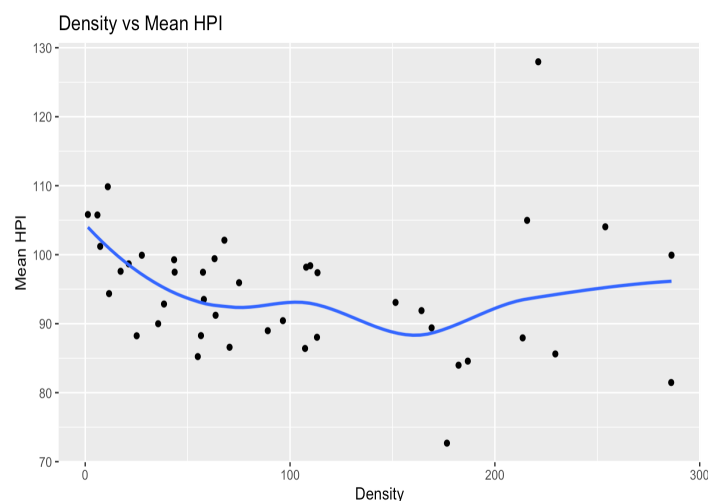
In f**igure 4.i** it looks like there is a trend in the sense that the HPI increases with density **but** it may be due to the presence of outliers. We remove the outliers using the Interquartile range. Outliers here are in the term of density values. We drop values that are greater than 1.5*IQR and below 1.5*IQR from the mean density.

**In figure 4.ii** we see the relationship between density and Mean HPI over the years after removing the outliers. Overall there seems to be no trend. The above plot captures the relationship across the US and maybe that could be the reason that a defined trend cannot be seen.

Here we again fit the loess model explained on page 3. This is due to the reason that the data could not be represented well by using a linear model.

Looking at it from 35000 ft, we plot the graph for HPI across years for present-day population density of all states.

We plot the House price index for all years against Present-day Population Density(**See figure 5**). We put the states in increasing order of their present-day population density. One startling insight is that the states with low population density recovered pretty well from the Great Depression. While the states with high population density never really recovered from the depression with their HPI values loitering around the same values as the pre-depression era even today with the exception of the District of Columbia. Apart from that, there is no clear & obvious trend in the changes of house price index with respect to present-day population density. We see that with increasing present-day population density, there is a similar trend for all states of the change in house price index across the years.

Figure 5

Figure 5. Changes in House Price with time against current Population Density
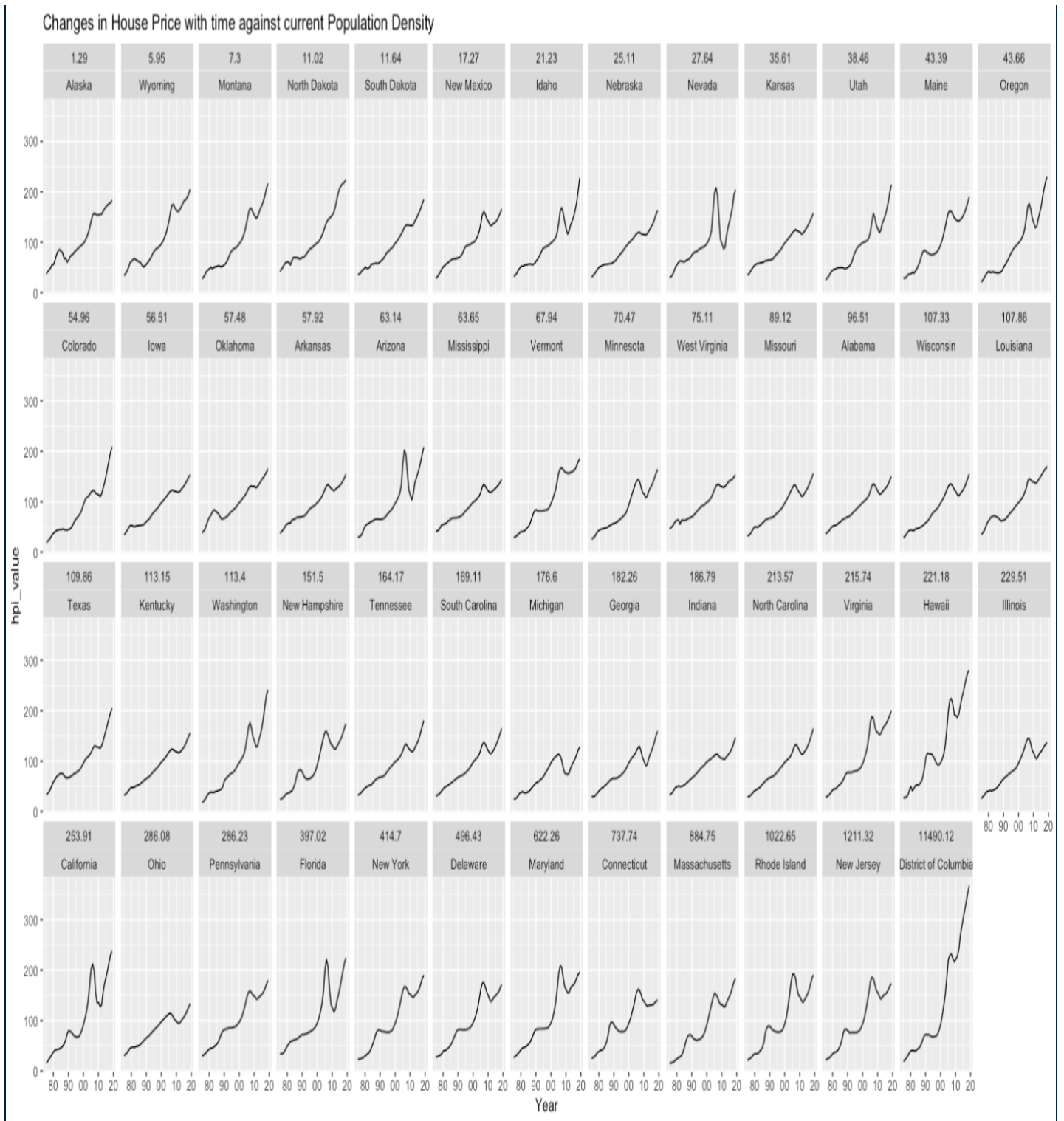
Now we see the trend region-wise(Figure 6) to see how population density affects house prices.

In the graph(Figure 6) we plot a linear model to observe the trend between density and house prices across regions.

Midwest region states: The fitted line suggests there is a negative correlation between house prices and density. The house prices decrease with the increase in density.

Northeast region states: There seems to be no clear trend between density and house prices.

South region states: There seems to be no clear trend between density and house prices.

West region states: The fitted line suggests there is a positive correlation between house prices and density. The house prices increase with the increase in density

5

FIGURE 6

Region Wise relation between HPI and density
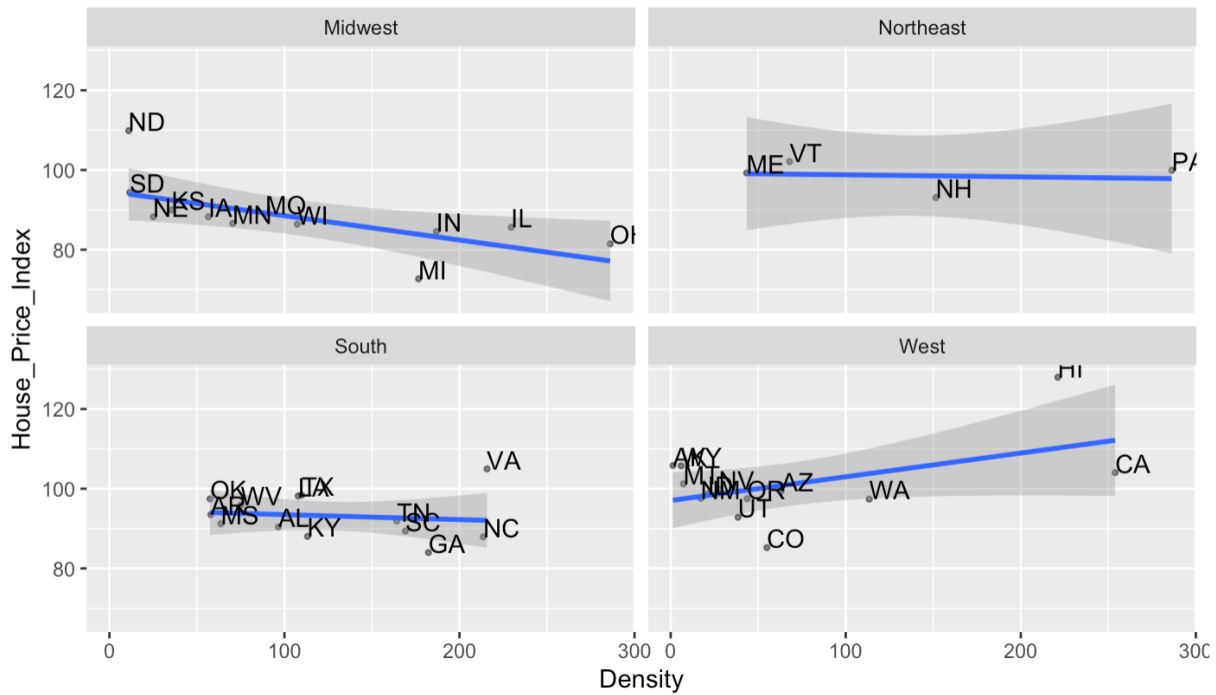
CHANGES IN POPULATION AND CHANGES IN HOUSE PRICES

We observed the relationship between the population density and house price and inferred that there was a relation between the two. Now, we see if there is a relationship between changes in population and changes in house prices. For this we observe data over 3 time periods: 1990 to 2000, 2000 to 2010, and 2010 to the present.

FIGURE 7

Changes in House prices with Population for all states across the 3 decades
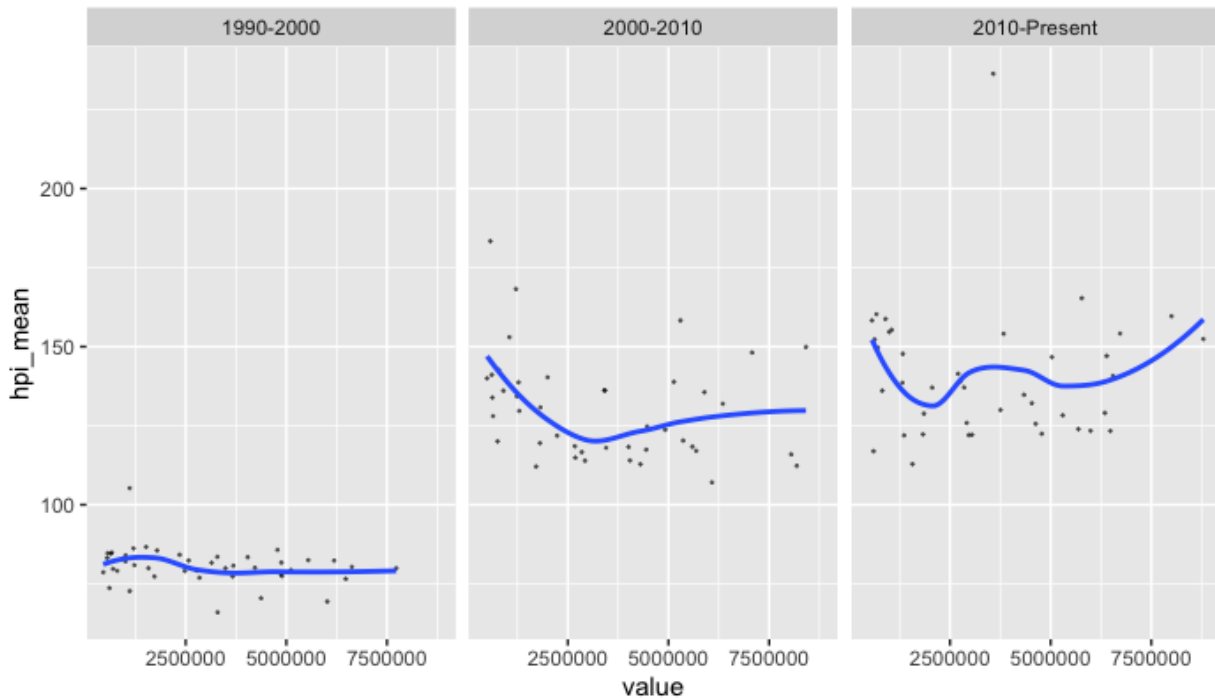
Here we again fit the loess model explained on page 3. This is due to the reason that the data could not be represented well by using a linear model.

1990-2000:

In the given decade 1990-2000, there doesn't seem to be any relationship between the house prices and the population (Figure 8) as the value of HPI is consistent even for the states with large populations. All the states have a value of HPI less than 100 except one state, which has the value of HPI marginally over 100.

2000-2010:

In the decade 2000-2010 also, there doesn't seem to be a concrete relationship as the HPI values (Figure 8) have increased over the decade, but this increment is not supported by the increase in the population. We can see that the population has remained almost the same as the previous decade, but the HPI values have increased significantly from around 50 to around 125.
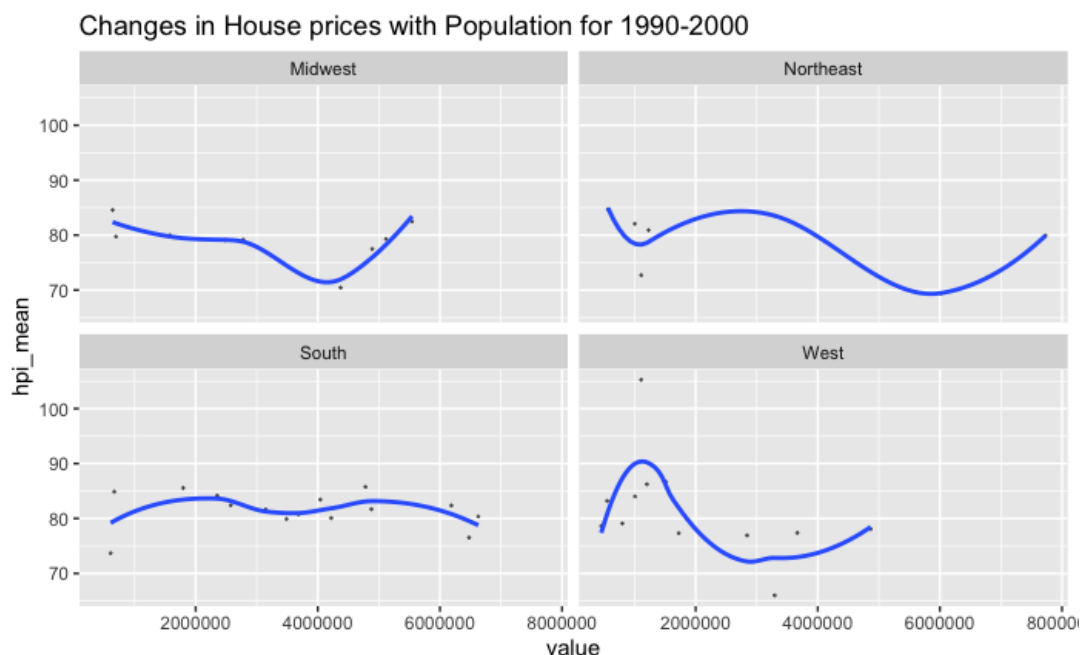
2010-Present:

Even in the last decade, it can be observed that there is no such relationship between population and the HPI values for all the states. Most of the states have the same population over this decade as well and the HPI values have also remained almost the same as the previous decade. Although, there seems to be some weird outliers in which the HPI value is almost 300 even for a relatively less population.

Plotting the population by regions:

1990-2000

For the decade 1990-2000, it can be seen that the HPI mean values have not increased over the increase in population values for every region.

FIGURE 8

Changes in House prices with Population for 1990-2000

2000-2010

For the decade 2000-2010, the relationship seems to follow a similar trend as compared to the previous decade as the HPI value seems to be independent of population values for all the regions.

FIGURE 9

Changes in House prices with Population for 2000-2010

2010-2020: For the 2010-Present decade, the house prices again follows a constant trend, with one or two outliers for the Northeast region.

FIGURE 10

Changes in House prices with Population for 2010-2020
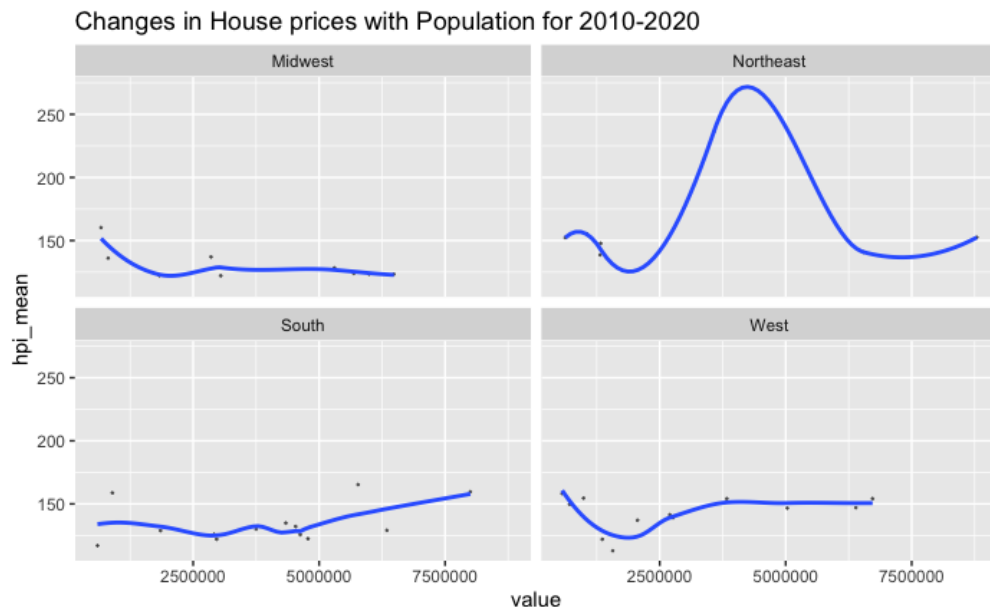
CONCLUSION

House Prices over time: The house prices have increased over time specifically from 1975 - 2019. I wouldn't say this inference is a surprise. But the house prices did go down drastically between 2007 and 2010 due to the great depression. The biggest increase in the house prices occurred in District of Columbia(6.11 %), the smallest increase in HPI was in West Virginia(2.51%). All regions of the United states follow a similar basic pattern and had a similar drop in 2007.

Population density and changes in house prices: The changes in house prices with population density can be best explained by region wise segregation.Outliers in density were removed to get better inference. The most obvious trend could be seen in the midwest and west.

Changes in the population and changes in the house: There can be seen a slight increase in the house prices in the decades 2000-2010 with a marginal increment in population . Although, there doesn't seem to be a fixed relationship between population and house prices over the three decades(for both states and regions).

8