EDA Project

Team 3 (Ojaas H, Vijeet S, Meet V)

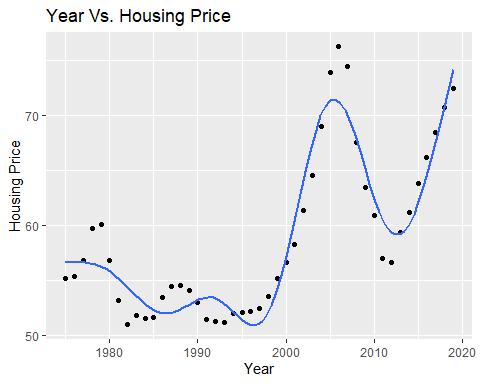
2/20/2020

## Introduction:

House prices and Population growth in the United States is a major topic which spans from the fields of census, economics to as far as determining the standard of living of the population in general. The main question to be answered here is whether the house prices have changed over the years and how is the change related to changes to population if any relation is found. This question is a complex question and needs to be answered in simple boiled down parts which can be assessed independently to get conclusions. The various sub questions are described in detail below.

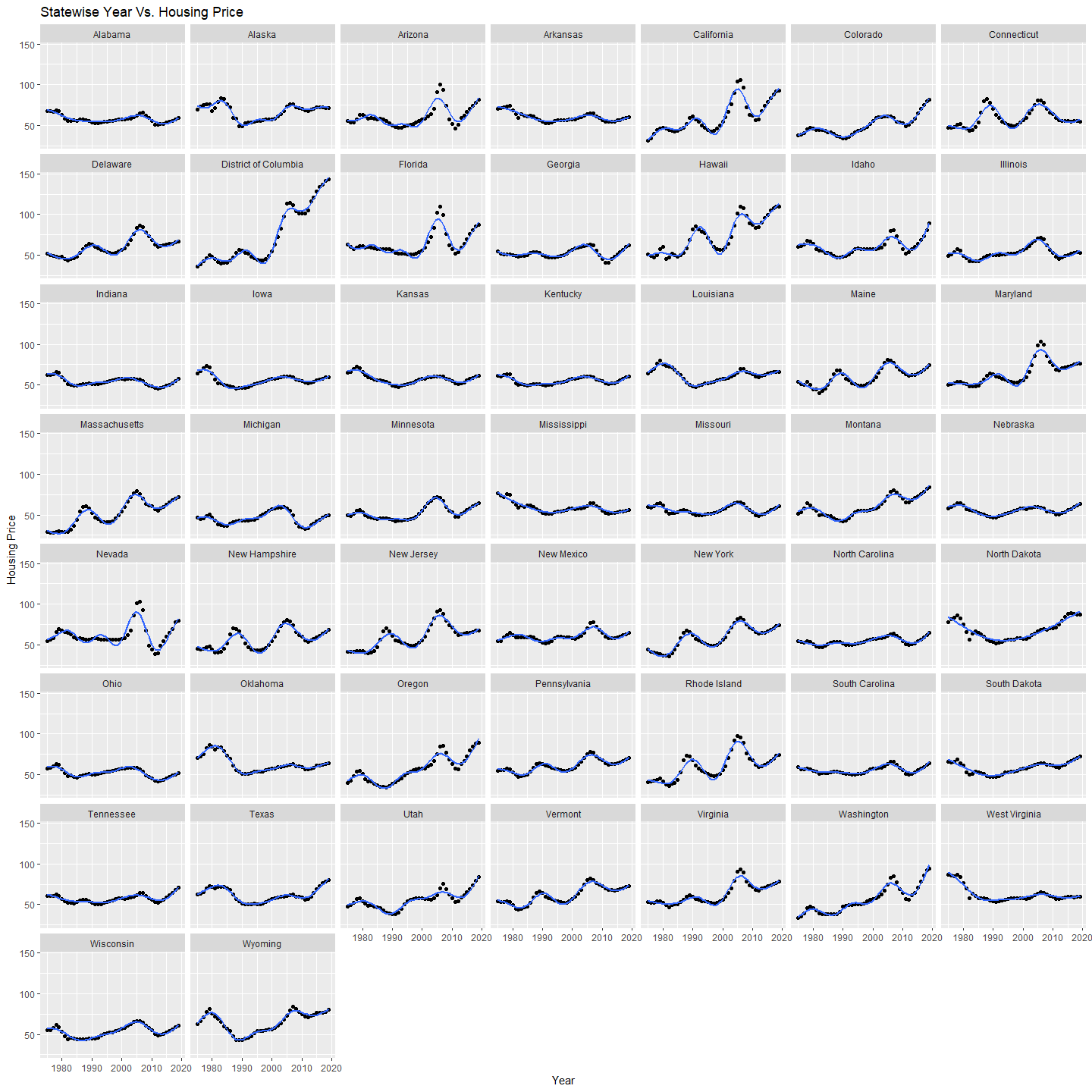
## Question 1: House Prices Over Time

## Change in House Prices US (CPI adjusted)

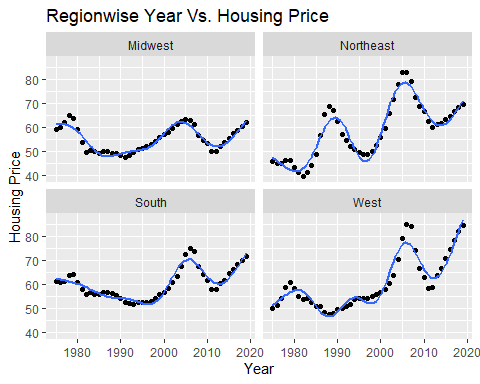


Year vs House Price: In this case, we have used a GAM model to fit the data. Gam is used to fit a non-parametric model to the data. GAM works very well when the factors to be compared is not particularly linearly related to each other. This can sometimes lead to a loss if interpretability, but in this case, we can easily interpret the change in house prices over the years. The graph clearly shows that there are fluctuations in Housing Prices over the time period.  The prices for the houses over the years have been adjusted according to the CPI Index (inflation), hence the effect of market conditions and currency changes has been taken into consideration before plotting the graph. There is a steep increase in the trend for increase in house prices after 1995.  The graph also shows that there are increased fluctuations in the 2000’s and forward. This effect can be attributed to the change in market dynamics and effects of globalization.

## Changes in House Price Varied by state



Different states showcase different relationships with house prices. Although each state obviously has a distinct curve that depicts relation between house prices and year, the underlying structure remains the same (showing an overall increase in price over the years). The district of Columbia has seen the biggest increase and Nevada has seen the maximum decrease in house prices. One state that can be equated as an outlier is West Virginia since its overall trend depicts a decrease in house price instead of increase, over the time span.



As seen region-wise, all the regions depict a similar trend of increase in house prices over the years especially in the later 2000s except Midwest which shows an overall flat trend as compared to all other regions. However, talking about absolute prices each region has its own numbers with West being the priciest of all, closely followed by Northeast. Midwest and South show quite identical trends though Midwest has an overall negligible change in house prices over the span of years.

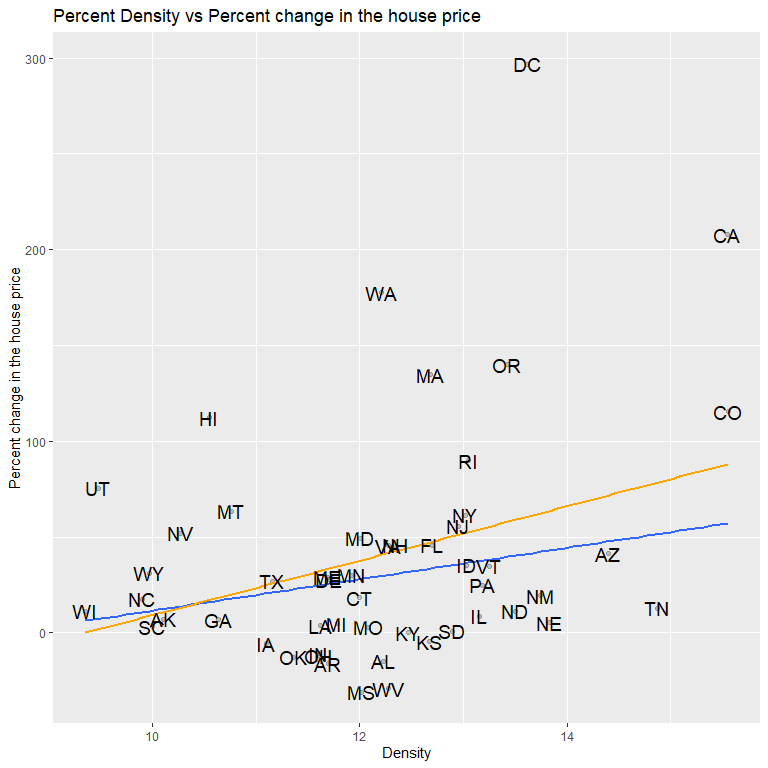
## Question 2

## To install your API key for use in future sessions, run this function with `install = TRUE`.

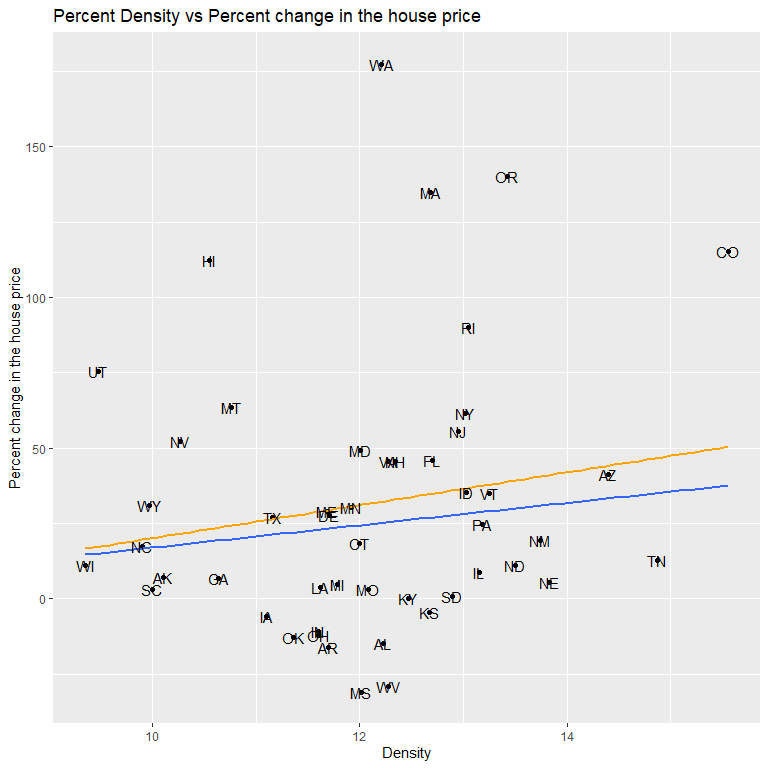
## Getting data from the 1990 decennial Census

## Getting data from the 2000 decennial Census

## Getting data from the 2010 decennial Census

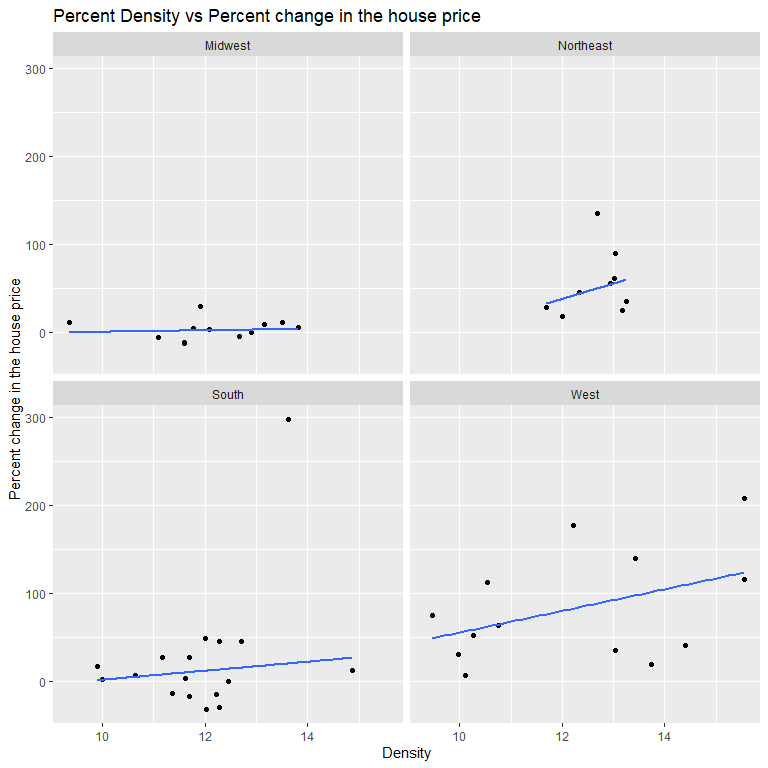


## Density vs Change in House prices after removing Outliers



With increase in density there is an increase in percent change in house prices. There are 2 outliers of DC and California which have been downweighed. This is done so because these 2 states do not follow the trend and change in house prices is extremely high. After this process the hypothesis is bolstered that with increase in density there is an increase in percent change in house prices. The removal of outliers considerably reduces the change in house prices.

## Change in House Prices vs Population Density (Region Wise)



Midwest shows no change in house price. Rest all three regions show increase in “percent changes in house price”. Though as previously observed South region shows a similar trend to Midwest which slight change in “change in house prices”.

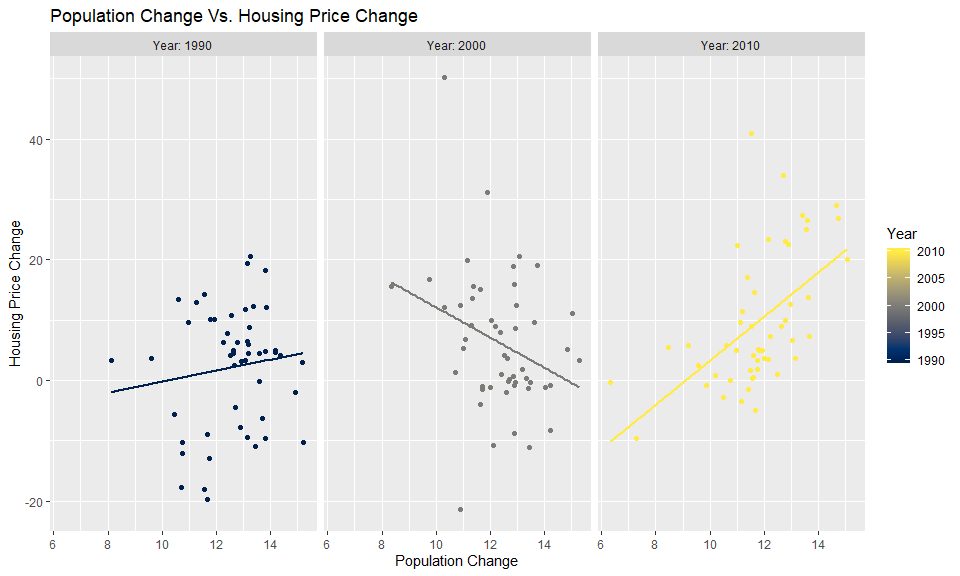
## Question 3)

## Getting data from the 1990 decennial Census

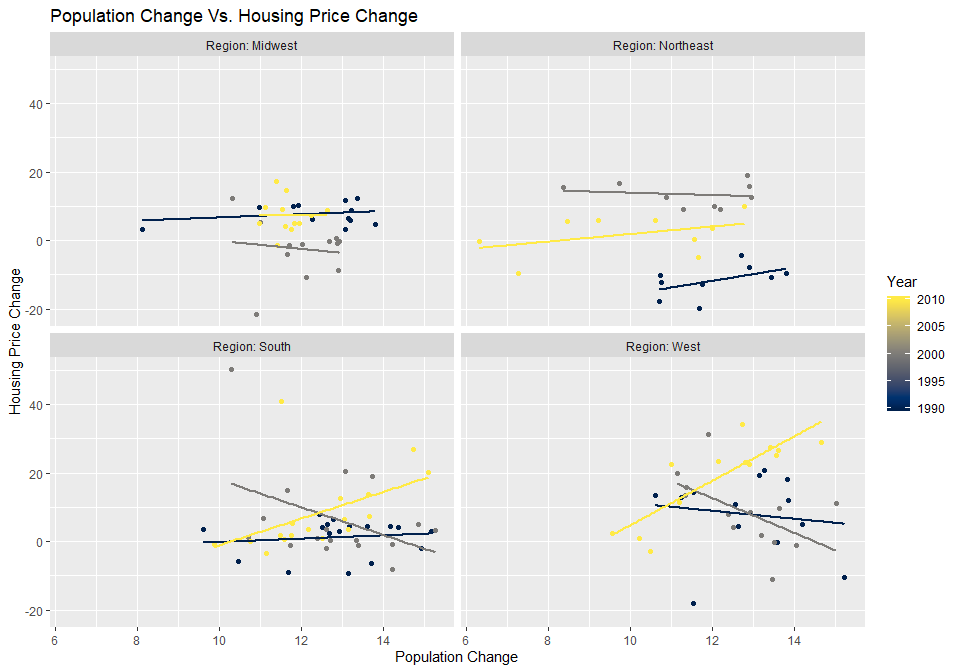
## Getting data from the 2000 decennial Census

## Getting data from the 2010 decennial Census

## Loading required package: viridisLite



A relationship can be observed between the two parameters of changes in population and changes in house prices. In the time period between 2010 and present the graph shows a positive trend which indicates a directly proportional relationship between the two aforementioned variables. In the time period of 2000-2010 there is a decrease in house price with respect to increase in change in population. There is a slight increase in change in house prices with respect to population change.



The region wise graphs compel us to introspect on the various trends between changes in population and changes in house prices. Starting with the year 2018 the graph shows an upward trend for all the four regions i.e. with increase in population change there is an increase in change in house prices. In the year range 2010s the population change is in inverse proportion with change in house prices. In 1990s there are no interpretable trends.

Conclusion:

The graph for house prices shows that there is an increasing trend for the house prices over the year. It is also a known fact that the population of various states has grown over the year. Keeping the above constraints in mind it is expected to observe a positive linear relationship between the changes in the population over the year with respect to the changes in house prices. The house prices here are adjusted to the CPI because we need to take inflation into account while dealing with the time series analysis of any financial data. Hence, on observing the data and knowing the logical trends and the graphs observed from the data plotted above, the natural conclusion that Changes in house prices with respect to change in Population shows a positive relationship can be derived.