A MODERN ANALYSIS OF HOME PRICE INDEXES IN THE USA

BY *Dieudonne Ouedraogo*

SUPERVISED BY **Prof Andrew Hernandez**

1.Introduction

The HPI is a broad measure of the movement of single-family house prices. The HPI is a weighted, repeat-sales index, meaning that it measures average price changes in repeat sales or refinancing on the same properties. This information is obtained by reviewing repeat mortgage transactions on single-family properties whose mortgages have been purchased or securitized by Fannie Mae or Freddie Mac since January 1975.

The HPI serves as a timely, accurate indicator of house price trends at various geographic levels. Because of the breadth of the sample, it provides more information than is available in other house price indexes. It also provides housing economists with an improved analytical tool that is useful for estimating changes in the rates of mortgage

defaults, prepayments and housing affordability in specific geographic areas.

The HPI includes house price figures for the nine Census Bureau divisions, for the 50 states and the District of Columbia, and for Metropolitan Statistical Areas (MSAs) and Divisions

In this paper I study how housing price in each US state as well as US major cities change during the years. I explore the difference in HPI change among west coast, east coast and mid west as well.I analyzed House Price Index in all 50 US states along with Washington D.C and 25 major cities and further developed a shiny app for interactive visualization.

2.Background and data

The data for my study are provided mainly by the Federal Housing Financial Agency and Bureau of Economy Analysis

*US housing price index summary
http://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_PO_summary.xls

*US housing price index by state
http://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_PO_state.txt

*US housing price index for metropolitan areas
http://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_PO_metro.txt

*Data for the largest cities in USA

http://travel.forumsee.com/a/m/s/p12-22806-0310459--metropolitan-gdp-billions.html
www.city-data.com

The goal here is to do a complete study of the House Price index by states, by division and to explore the variation for the largest cities in relationship with the annual income.

A summary reveal that the House Price Index (Purchase Only Index) for the whole country shows that the maximum of HPI index is 225.2 and the median is 160.9. (where index for 1991-Quarter1 is set to 100) from 1991 until 2016

3. Analysis of HPI for all states

Average HPI for each year by US Division
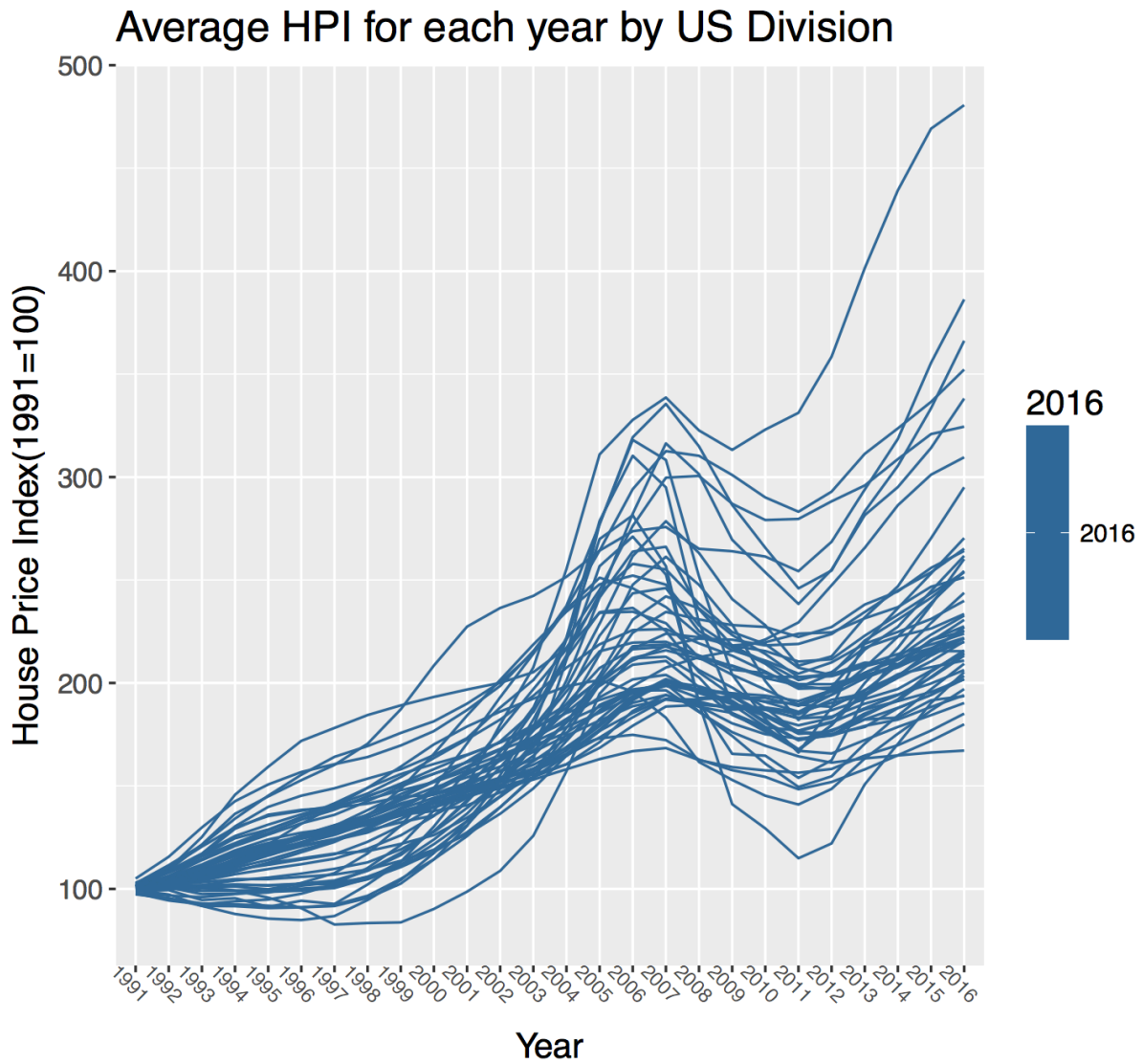
I used package "lubridate" to deal with Dates. Year and Quarter were switched into DATE. Each quarter is represented by the first day of the beginning of the month in the quarter, eg.2007-01-01 represents the first quarter, 2007-04-01 represents the second quarter, 2007-07-01 represents the third quarter, and 2007-10-01 represents the fourth quarter. Then I averaged over 4 quarters of each year and generated a parallel coordinate plot according to years as well as making a time series plot for all quarters.

Next we can make a parallel Coordinate Plot of Purchase Only Index for each state in each year, grouped by US divisions and regional classes. From the result we can observe that HPI rises slowly from 1991 to 2000, and then increases exponentially since 2000 and reaches peak in 2008. After the year 2008, HPI drastically fell down and was fairly low in 2011-2012. However, we could also notice that from the year 2013, it seems that the

house market is thriving and perhaps it is turning over. This is a positive sign for the turn around of economy, although we need more data to determine whether the house market is recovering for sure. Moreover, the plot also shows that DC has not been affected by economy crisis significantly during 2007-2008, The HPI fell down a little bit and then went up immediately. HPI for DC is extremely high during these years compared to other regions in US based on the fact that the HPI is always higher among all other areas since 2005.

4.Average HPI for each year by US Division

## Average HPI for each year by US Division
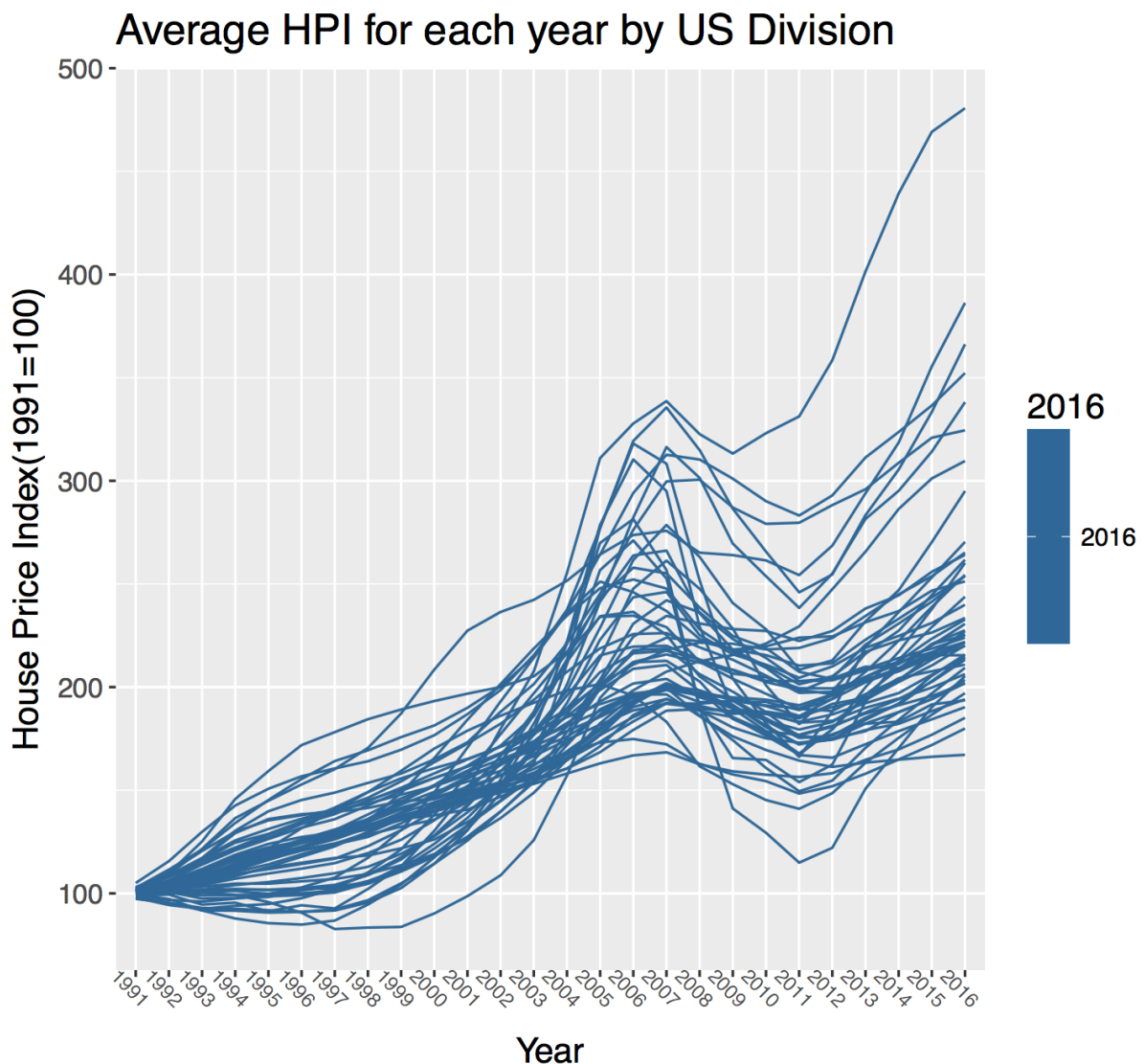
I grouped the time series plot by different clustering criteria after dividing the US states roughly into 5 regions, we could easily find that on the West part has higher HPI index than other areas despite the fact that DC stands out.

5. Average HPI for each year by US Regional Class

Average HPI for each year by US Regional Class
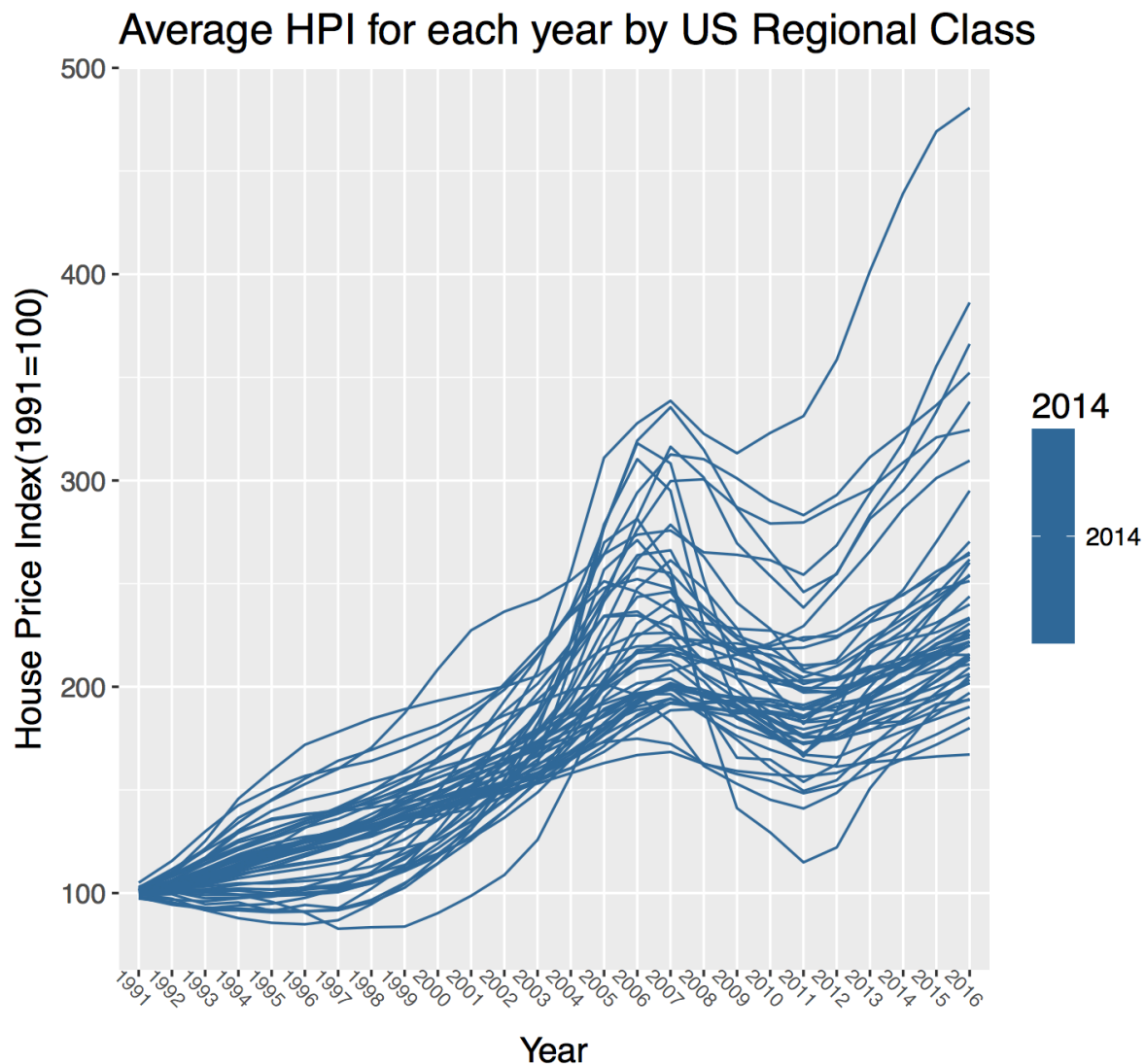
Finally, I also made a parallel coordinate plot for all four quarters in the period of 1991-2016 by division. DC still has the highest HPI and mountain and pacific areas have relatively high HPI during these years. Besides, different quarters do not make much difference within a year therefore mean value of HPI for each year is enough for analyzing the trend.

5.All Years and Quarters

All Years and Quarters
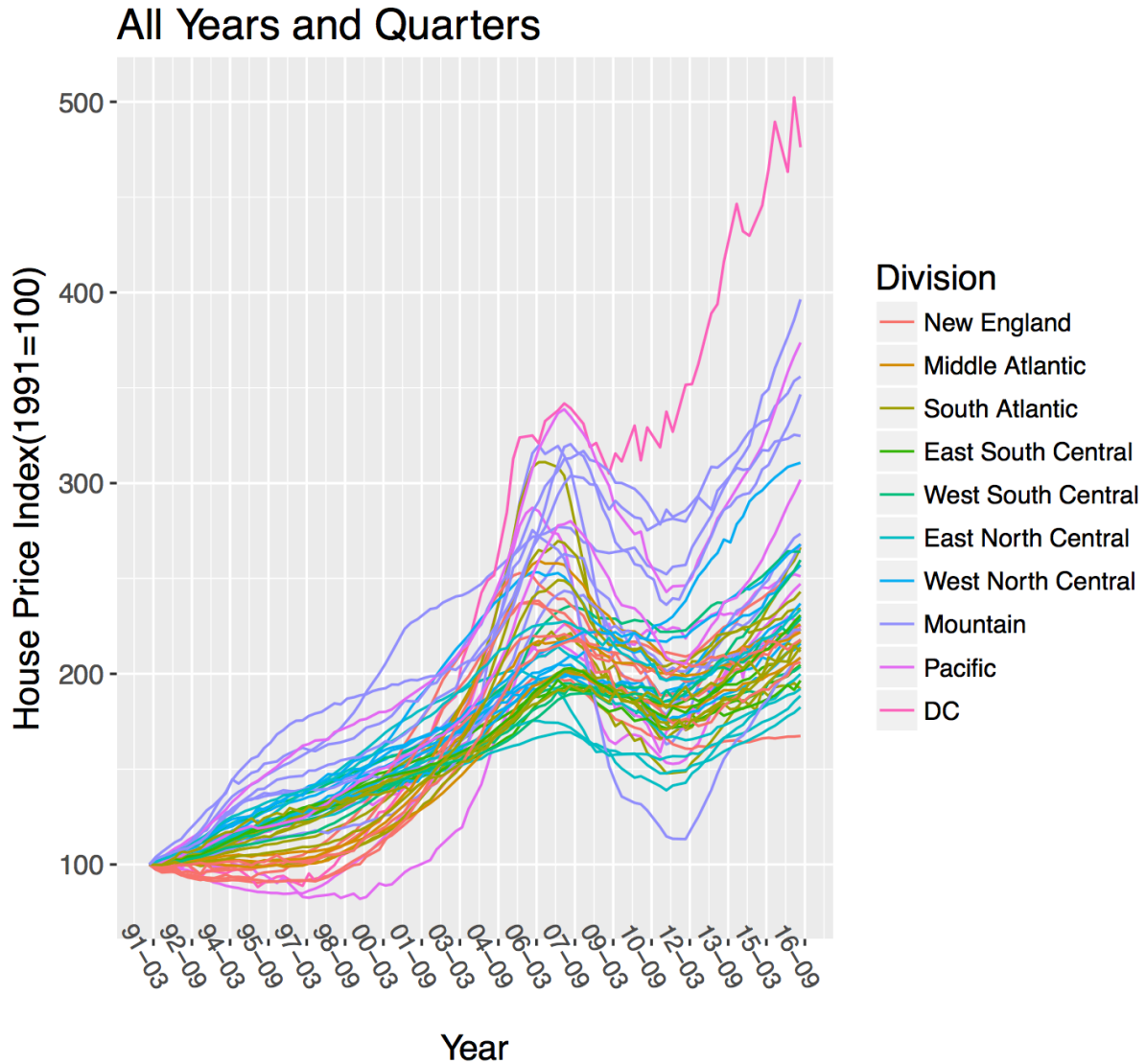
Above is the plot all quarters since 1991 ,by division ,DC standout as well

6. *Analyzing the percentage changes of HPI for all states*

I then analyzed percentage change of HPI (Purchase Only Index) over previous quarter, previous year, past 5 years and since the first quarter of 1991. D.C and USA result is also included in the analysis. Period ended is 2016-Q3. In dataset US-HPI, since the region "USA" has leading and trailing whitespaces, we need to trim it off to combine US HPI data with US map data.

We can then sort the HPI changes grouped by division over, for example, the past One-year.

The sorted result shows that, Over the past one year from Q3 of 2015 to Q3 of 2016, the HPI change in Pacific area increases significantly (past One year). The second highest region is Mountain area. These two regions carry higher NPI increase than that of the whole country. States in Pacific regions are: Alaska, California, Hawaii, Oregon, Washington and states in mountain regions are: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah and Wyoming. This makes sense since we know that housing prices in California and Alaska are always expensive as well as in its surrounding areas such as Arizona and Nevada. This also indicates that over the past year, housing market has come back to life originated from the west coast and pacific areas.

Division	OneQtr	OneYr	FiveYr	allYr
Pacific	0.8744	6.54	39.52	179.6
Mountain	1.366	6.238	38.98	199.6
East North Central	1.532	5.694	22.84	99.35
East South Central	1.696	5.39	19.79	112.4
South Atlantic	0.9193	5.094	25.62	124.5
West North Central	1.267	5.063	24.95	148.8
West South Central	1.185	4.176	23.99	139.1
New England	0.7388	3.469	13.9	115.9
Middle Atlantic	0.977	3.129	11.33	118.8
DC	-5.21	-2.729	45.64	376.2

Percentage change

Take a look at HPI percentage changes for all divisions in the given time span: change over previous quarter, over previous year, over past five years and over all years since 1991-Q1. We found that over the past five years, HPI change is very small compared to changes over all years. This could be explained by the fact that the US has gone through an economy turndown since 2008. However, for the past 20 years, the Housing Price Index has increased by almost 100 percent for around half of the US states and HPI surge in DC has even increased by 375 percent.
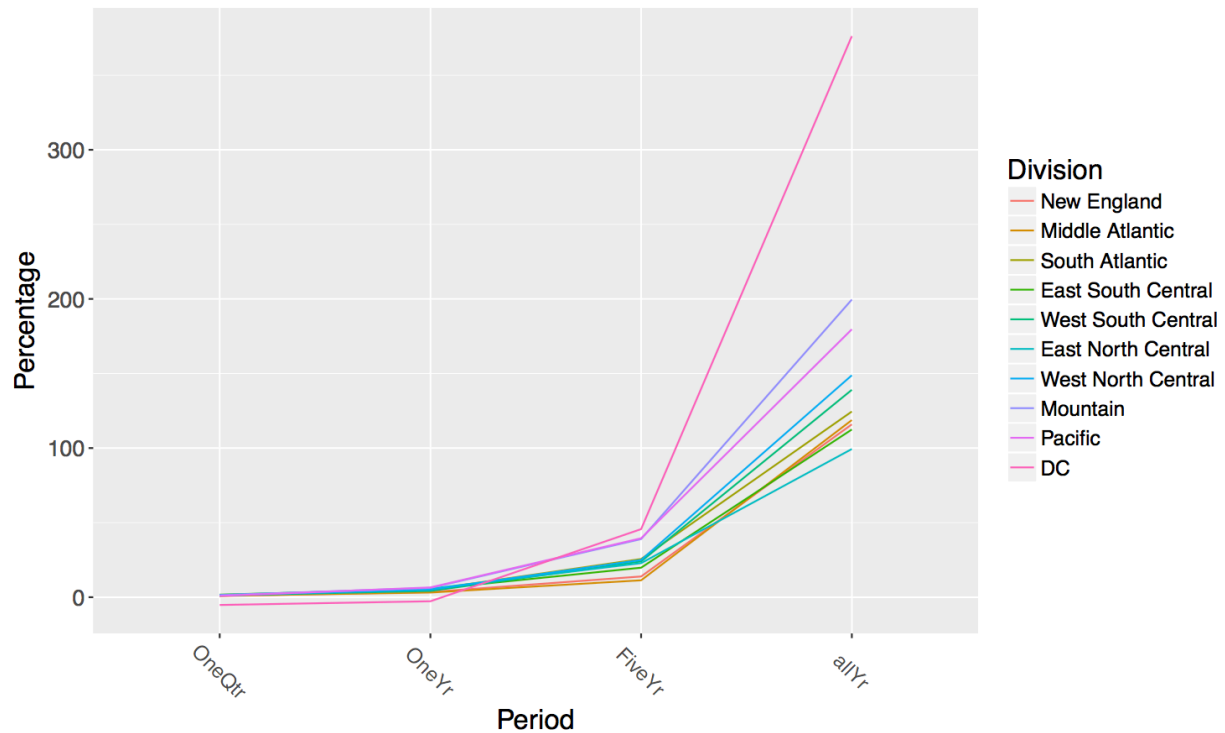
We could also use bar plot to visualize the HPI inrease during different period of time. The upper left panel in the following result shows that for the past quarter, from 2016-Q2 to 2016-Q3, HPI decrease of DC is significant while HPI of mountain and pacific area has increased most. This is unusual since HPI of DC is always increasing and is always higher than other part of US states. The upper right panel is HPI changes over the past year for all divisions. Surprisingly, we found that the HPI increase is at least 2.5% during

2015-2016. This shows that housing market is starting to turning around in most areas in US. In Five-Year range in the lower left panel, most parts of US have low HPI percentage change (or increase) since the country has been just recovered from a crisis happened years ago and HPI was coming back slowly. From the figure I can conclude that the housing in mountain and pacific areas has been recovering most in recent years.
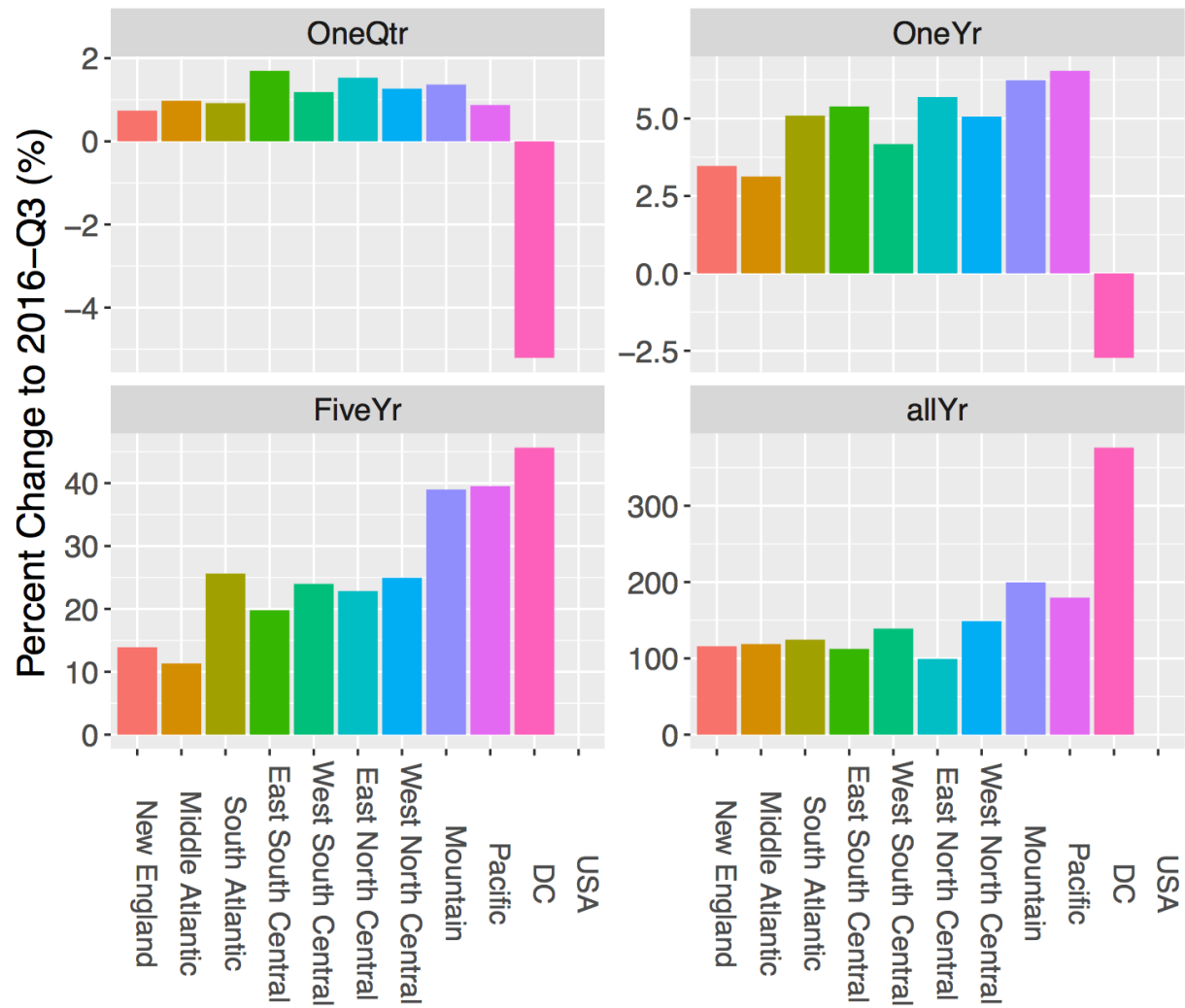
7. Visualizing HPI changes of 50 US states and DC

House Price Index in 2013–Q4

Purchase Only Index 200 240 280
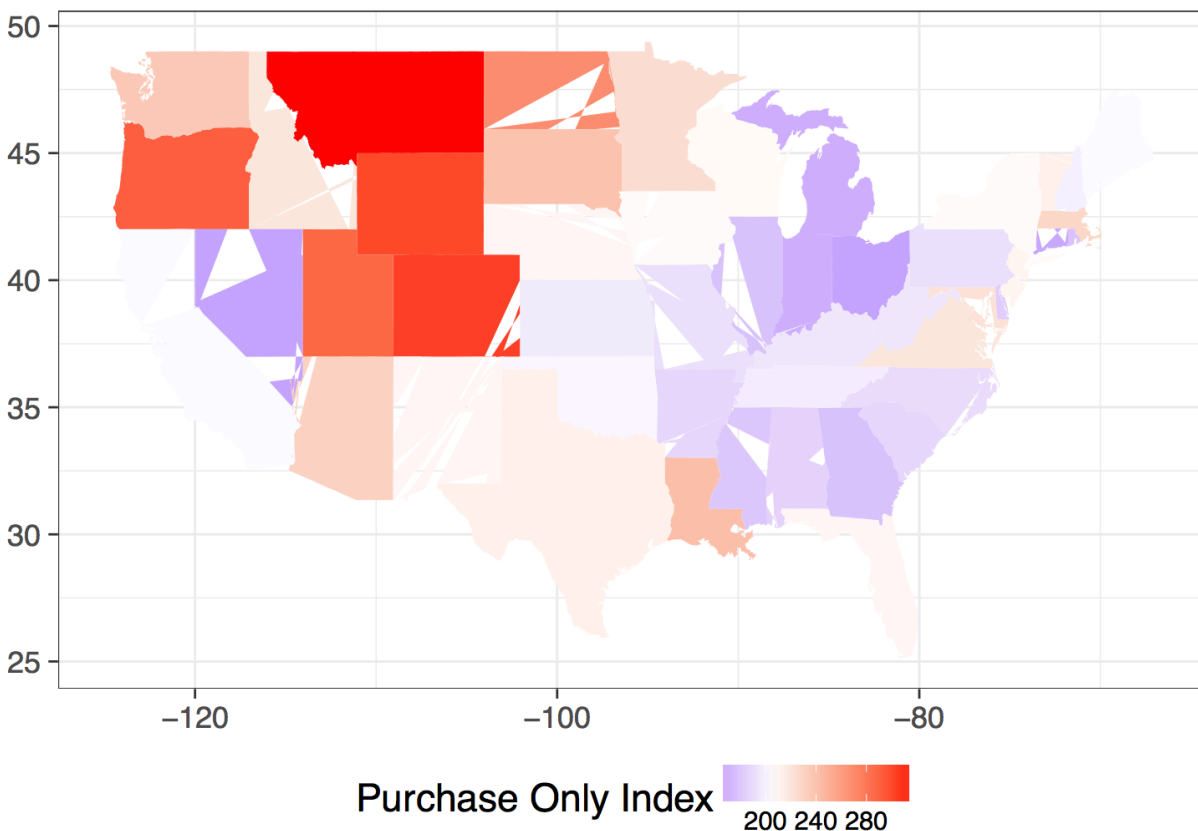
The HPI absolute value and HPI change for each state could be projected onto US map for better visualization. The following figure is a plot of HPI value for the most recent period 2013-Q3 for each US state on US Map. On the plot, we use HPI index of USA as the value of midpoint (white color on the map). Red represents high value and blue represents low value. In the most recent months, HPI of Montana, Wyoming, Colorado and Oregon have the highest HPI while Indiana, Michigan and Ohio have the lowest.

Next, we also made plots of HPI changes for each state in US over a certain period of time and then project the percentage change on US map. The mapping code is written into a function and we can choose time span of interest, such as one quarter, one year, five years, etc.

Change in one quarter

The figure below is HPI change over one quarter during 2013-Q3 and 2013-Q4. Nevada, Arizona and Hawaii are the top three regions that have the largest HPI change in this period of time, with more than 3% of HPI increase. HPI change for overall USA is

+1.20% and there are 18 states that has higher HPI change than the average HPI change for USA.

Change in One year

As for one-year duration for the whole year 2013, HPI change of Nevada, California and Arizona rank the top three among all other states. The HPI increase is 24.3%, 19.5% and 15.2% respectively.

Change in five years

For HPI change over a long span 5 years since 2008, Washington D.C, North Dakota and California rank top three, with HPI increase of 27.8%, 27.7% and 19.4%. The amount is surprisingly almost the same as one-year increase of HPI in 2013 for West coast.

During 20 years since 1991, HPI change increase most in Washington D.C area, followed by Montana and Colorado.

8.Housing Price Index and Per Capita Personal Income for 25 Metropolitan Cities

Next, relationship of HPI change and Per Capita personal income is analyzed. The first thing to do is to download Per Capita Personal Income for 25 metropolitan cities on Bureau of Economic Analysis. (http: //www.bea.gov) The exact steps to download the data of interest are described in section2(4). Income change is preferred and we are only interested in the past five years and 25 biggest cities as well. One thing to note is that on this website, the latest income change is from 2011-2012. Since defining of regions or areas is quite messy on different websites, I matched the city and counties in the downloaded dataset with information of 25 biggest cities. Then we need to preprocess the us. Cities data by adding D.C entry. Here I used Arlington, VA as the latitude and longitude of D.C since Arlington is quite close to D.C.

The following figure is a parallel coordinate plot of Personal Income Change during the past years. Since year 2007, the personal income changes first decreases to negative and then increases. During the period of 2008-2009, income change rate reaches lowest-below zero for all 25 biggest cities. However, during these years, personal income change rate speeds up and exceeds the change rate before economy turndown.

Per Capita Personal Income Change
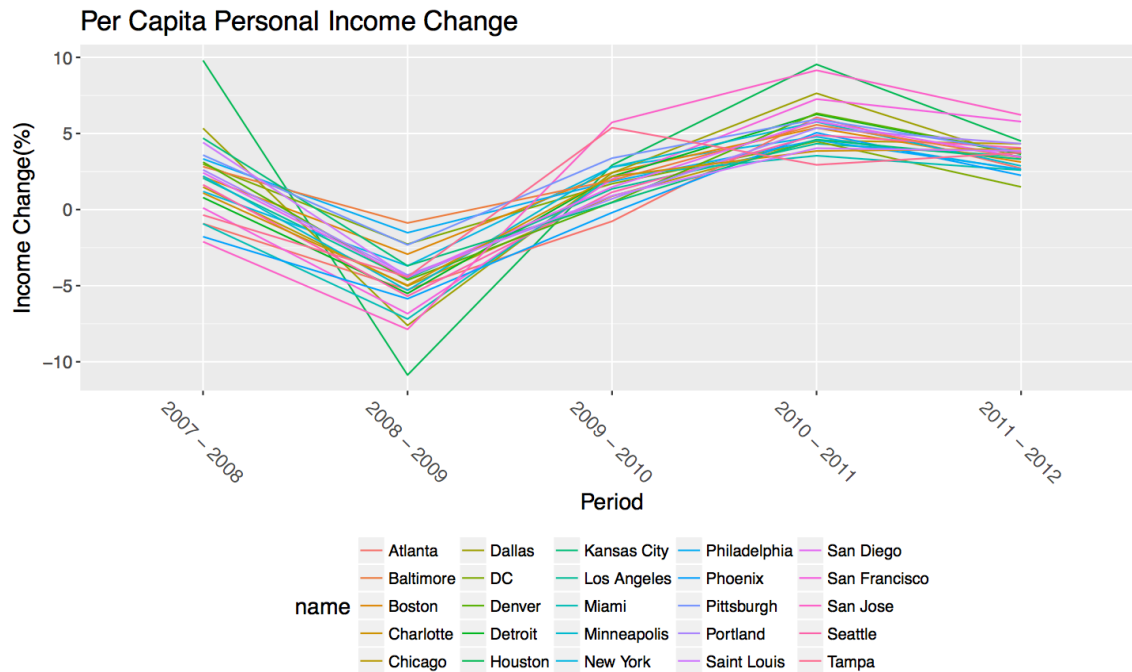
After loading the dataset, processed and combined them, we could map HPI change and Per Capita Income change on US map for a time period such as 2011-2012. The following figure shows percent change of HPI and income in year 2011-2012 shows the relationship of HPI and income changes in these large cities. Blue represents low HPI change and red represents high HPI change. The change in income over one year is represented by size of points. The bigger the points, the higher the increase of Per Capita Personal Income. We can tell that Personal income increases most in west coast such as San Francisco and San Jose with relatively low HPI increase. The HPI increases most in Baltimore, MD which is near DC area with an increase of 26.6%.

A glimpse of HPI and income change during year 2007-2008 is shown below. Compared to last two years' boom, HPI decreased drastically in year 2007-2008 and personal income increase was quite low in most areas. Houston has the highest HPI decrease rate of -25.3% although its income increase is 9.80%, highest among other areas. Most areas have undergone an HPI decrease which marks the fell down of housing market despite that some popular areas such as San Diego and New York were not significantly affected by the economy turn down.

This analysis on HPI and income change was also turned into a vivid Shiny App that can show the relationship of HPI change and income change in major US cities during a certain period of time of interest. The following two screenshots were taken from my

9.Shiny App.

In this easy App, there are three tabs. The first tab Plot shows HPI and personal income changes in a given period of time. On the left panel, user can select the period of time they are interested in. The visualization can also be adjusted by hiding/showing legend and labels as well as changing size and transparency of points and label texts. The second tab Summary displays a summary of dataset for HPI and income during a selected period. We can know about min(max) of HPI and Income change. The third tab is a table view of the dataset, I used renderDataTable which is quite fancy to show results including sorting, change number of observations, search by keyword, etc.

House Price Index for US states and 25 Metropolitan Cities

House Price Index for US states and 25 Metropolitan Cities

Choose the period of year :

Year 2011-2012 ▼

☐ Hide Map Legends

Choose Size of Points:

6   15   30

Choose Transparency of Points:

0.1   0.5012   0.8

☐ Hide City Labels

Choose Size of Labels:

1   5   10

Choose Transparency of Labels:

0.1   0.5012   0.8

First, select a period of time to see Relationship of House Price Index and Per Capita Personal Income of Metropolitan Cities in US. You can change Visualization views by dragging slider bars below.

Plot   Summary   Table

Show table of dataset and sort cities by hpi ~ income11_12

10 ▼ records per page   Search:

city	state	income	hpi
Baltimore	MD	3.41	26.63964
Seattle	WA	4.05	17.43831
Minneapolis	MN	2.87	15.20790
Miami	FL	2.59	15.19715
Denver	CO	3.70	12.21241
Portland	OR	4.33	11.76664
San Diego	CA	3.44	11.74399
New York	NY	2.60	11.66906
Houston	TX	4.50	11.35180
San Francisco	CA	5.78	10.69690
city	state	income	hpi

Showing 1 to 10 of 25 entries   ← Previous   1   2   3   Next →

10. Discussion Conclusion and Future Research

This research project helps me to get a deeper knowledge and substantial practice in data technology skills. The skills applied in the project which I have learned in the MSDA program. After analyzing the datasets, I could conclude that during the past five years US economy has been turning around with a step-by-step rise in housing market. In this project, I have analyzed the HPI and HPI changes during different period of time as well as changes in per capita personal income at the meanwhile. I have decided to choose personal income rather than Gross Domestic(State) Product (mentioned in the proposal) in that personal income is more specific and better to reflect the purchase power of people in the big cities. In terms of HPI and HPI changes, Washington D.C stands out with highest HPI value as well as high HPI changes except for the 4th quarter in 2013. One year's HPI increase is quite much in most regions in US with Mountain and Pacific areas rank top. Compared to five years' period, HPI increase is fairly low below 10% in most areas except that D.C experienced an increase of 26%. This shows that US housing market has fell down since 2007-2008 and now it has just come back to life. In the coming 2016-2017 period, the housing price is expected to rising according to the trend on my plots

There are other aspects of the project that are not cover here due to the limit amount of time and my goal in the future is to cover and tackle those aspects I need to further investigate the statistical properties of repeat transactions house price indexes and there are many topics to be explored such as geographic and temporal aggregation, revision volatility, the use of appraisal values, sample selection, and comparisons with alternative methods.

10.repository and links

*github to my project

https://github.com/dieudo/DATA698

*Published Shiny app for my project

https://dieudonne.shinyapps.io/dieudonnecapstoneshinyapp/