Team 7:Huy Tu, Nabaruna Karmakar, Samuel Heroy

Problem Statement:

The onset of the 'Great Recession' was preceded by decades of growth in the housing market, with the pricing of homes in many cases tripling between 1999-2007. The crashing of this bubble led to the dropping of home values across the United States, which ultimately led to the great recession. Unemployment rates over the country reached a record high in 2009-2010, a drastic increase from 2008 as a direct impact of the real estate market crash.

We analyzed the trend in home value index, by looking at the "real-estate" dataset and the number of job postings across the country, using the "jobs" dataset. We also included demographic data for various cities and states in the U.S. in our analysis and quest to find a correlation between home value index and number of jobs created each year, between 2007-2016.

Non-Technical Summary:

We looked at the change in terms of job creation, depending on population in 2007 and 2016. The figures below show a geo-spatial visualization of the number of jobs created in 2007 and 2016. It can be seen that in 2007, the number of job creation was very low, even in large cities. While in 2016, after recovering from the recession, it seems that there was more job creation, in more cities across the U.S. and even the big cities showed more number of jobs. California seems to have been least impacted by the recession, while Chicago and New York started seeing an increase in the number of jobs after recovery.

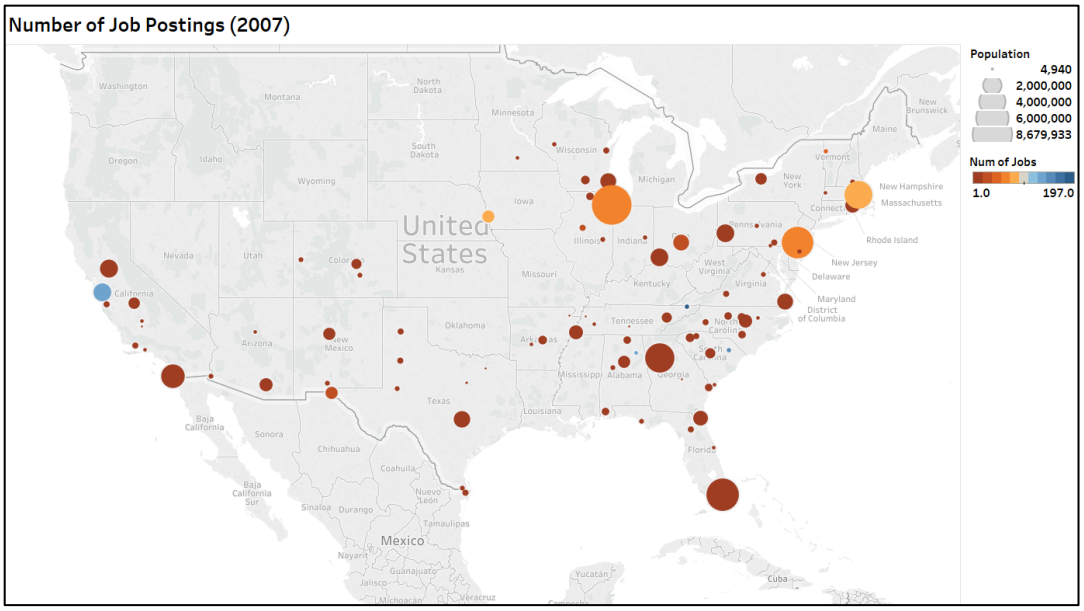

Figure 1: Number of Job Postings in 2007

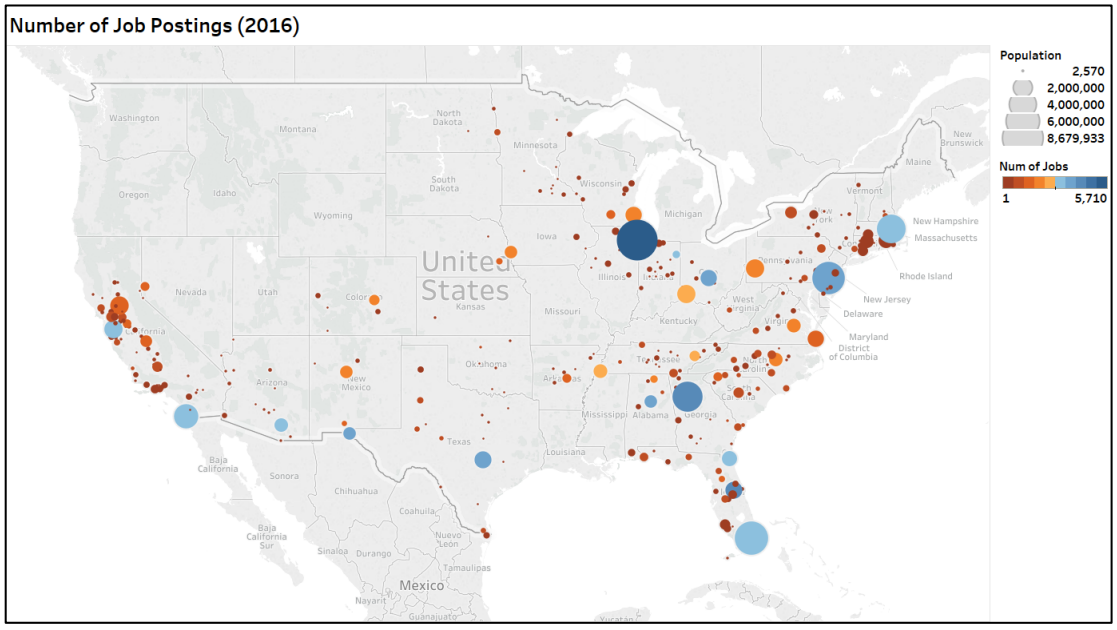

Figure 2: Number of Job Postings in 2016

Technical Summary:

The steps for our approach is displayed graphically in the figure below.

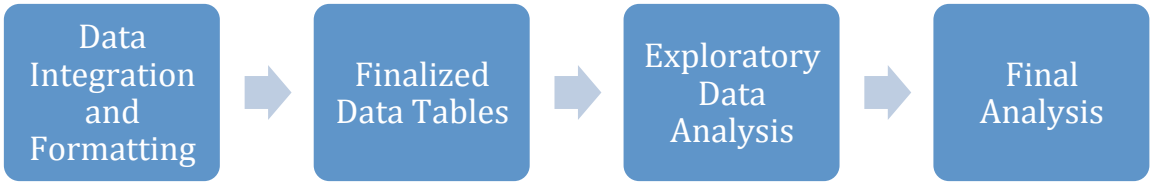

Figure 3: Steps of Data Analysis

1. Data Integration – We integrated data from various sources. First, we used the jobs postings data and came up with the total number of jobs posted each year, in each city and state. We then used the real estate data to integrate an average Home Value Index for every year in each city and state to the jobs data set. Lastly, we incorporated the city demographics data to the jobs and home value index data set to create out final table. A snapshot of the final table is displayed below.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Row_num	City	State	Year	SizeRank	Home Value Index	Num of Jobs	population	households	median_household_income	mean_household_income	
2	1	Aberdeen	MS	2015	4382	73233.33333	3	4483	1716	19050	19050	
3	2	Aberdeen	MS	2016	4382	75533.33333	2	4483	1716	19050	19050	
4	3	Aberdeen	MS	2014	4382	75900	3	4483	1716	19050	19050	
5	4	Aberdeen	WA	2012	2026	98550	16	29367	12837	25232	25232	
6	5	Aberdeen	WA	2015	2026	96783.33333	87	29367	12837	25232	25232	
7	6	Aberdeen	WA	2011	2026	102916.6667	8	29367	12837	25232	25232	
8	7	Aberdeen	WA	2013	2026	96958.33333	56	29367	12837	25232	25232	
9	8	Aberdeen	WA	2010	2026	114683.3333	5	29367	12837	25232	25232	
10	9	Aberdeen	WA	2014	2026	93366.66667	69	29367	12837	25232	25232	
11	10	Aberdeen	WA	2016	2026	116483.3333	173	29367	12837	25232	25232	
12	11	Abernathy	TX	2012	6830	69366.66667	1	3277	1694	26024	26024	
13	12	Abernathy	TX	2016	6830	70466.66667	1	3277	1694	26024	26024	
14	13	Abilene	KS	2011	6775	93700	20	6787	3549	27360	27360	
15	14	Abilene	KS	2013	6775	92925	15	6787	3549	27360	27360	
16	15	Abilene	KS	2015	6775	93008.33333	13	6787	3549	27360	27360	
17	16	Abilene	KS	2014	6775	91433.33333	27	6787	3549	27360	27360	
18	17	Abilene	KS	2012	6775	95825	8	6787	3549	27360	27360	
19	18	Abilene	KS	2016	6775	92883.33333	21	6787	3549	27360	27360	
20	19	Adel	IA	2014	4513	188366.6667	1	3583	2055	34947	34947	
21	20	Adel	IA	2013	4513	181116.6667	2	3583	2055	34947	34947	
22	21	Akron	OH	2011	123	62808.33333	164	565839	299631	27631	27631	
23	22	Akron	OH	2012	123	60691.66667	225	565839	299631	27631	27631	

Figure 4: Finalized Data Table

2. Exploratory Data Analysis – To get a sense of our hypothesis, we looked at the trend in Home Value Index and Unemployment Rate across the USA between 2007 and 2016. The figures below show the trend in average home value index and unemployment rate between 2007 and 2016. A large dip in the average home value index is noted between 2008-2011 because of the recession in Figure 1, while Figure 2 shows a steady increase in unemployment during that time. Our data integration was performed in Tableu, and our data analysis was performed in R.

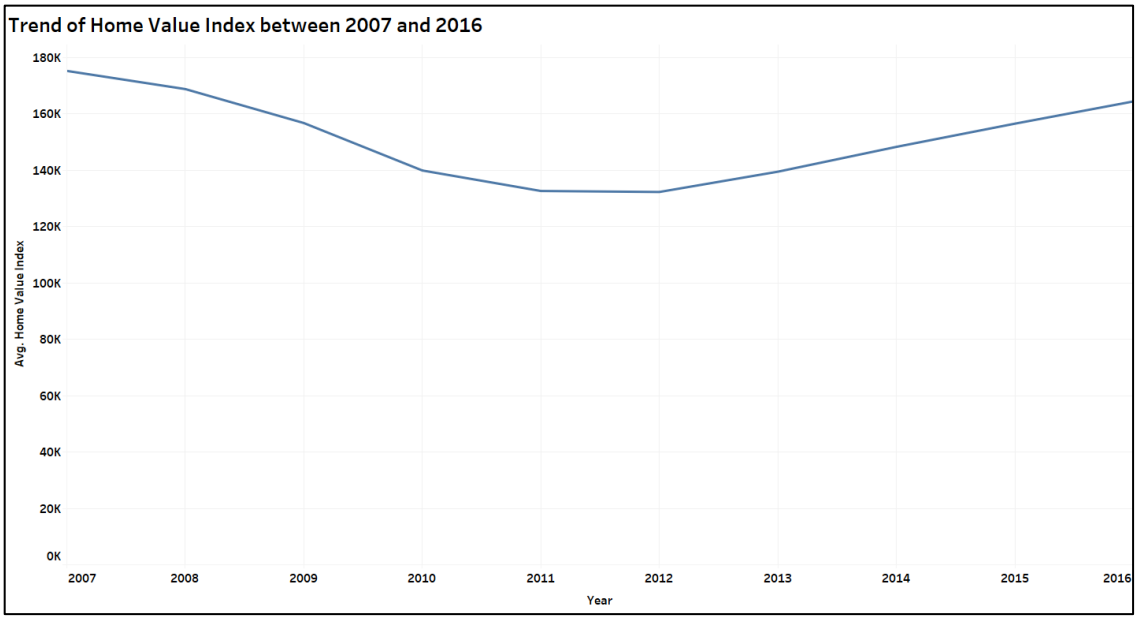

Figure 5: The change in the Home Value Index between 2007 and 2016

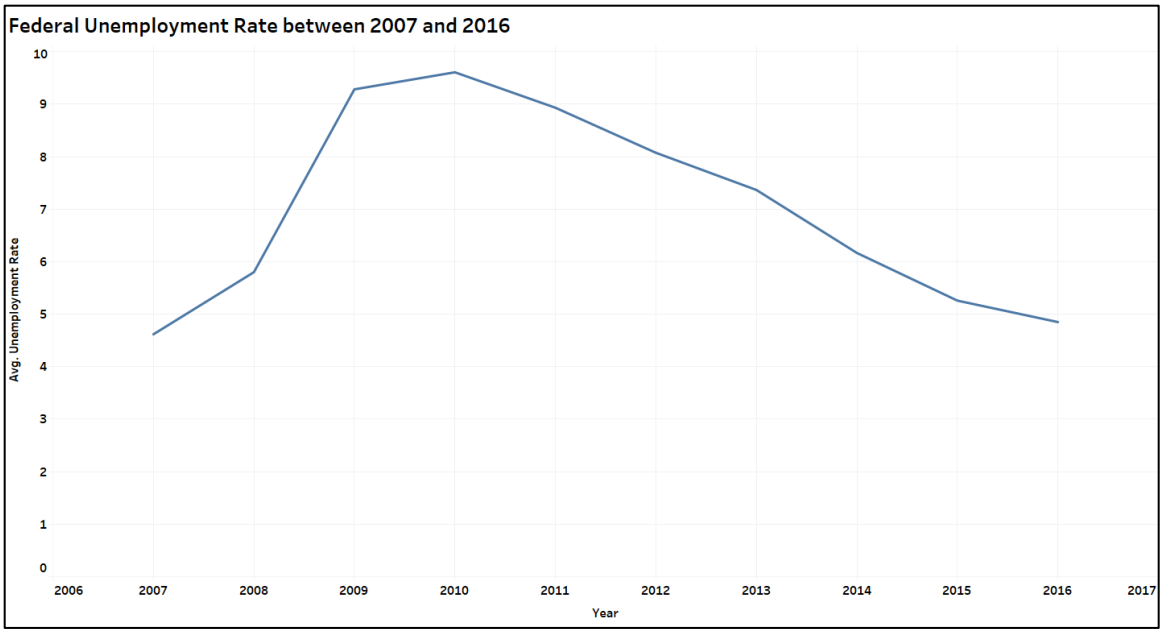

Figure 4: The change in Average Unemployment Rate between 2007 and 2016

In Figures 3-4, we see a drop in the home value index in the critical years of the housing crisis and a corresponding spike in unemployment for the years of the financial crisis.

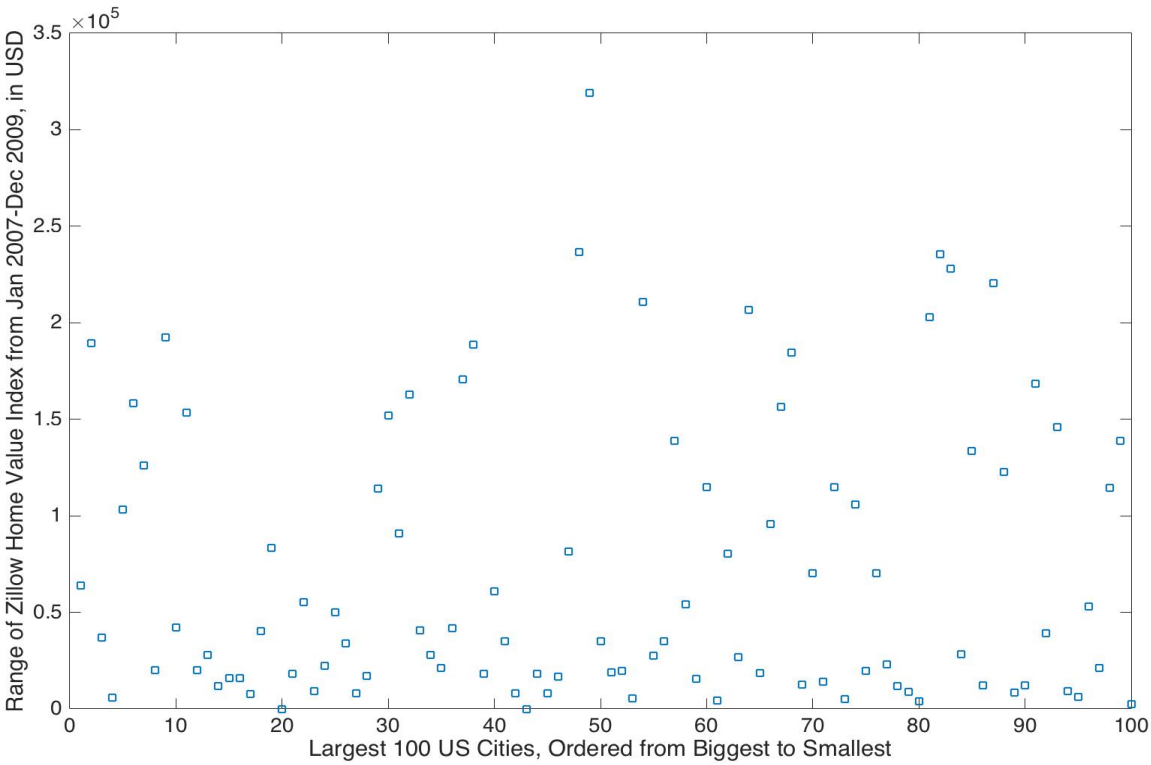

Figure 5: Zillow Home Value Index for largest 100 Cities in the US in terms of population, left being the smallest.

In Figure 5, we see that there is no relationship between population and the range of home values (largest minus smallest between 2007-2009).

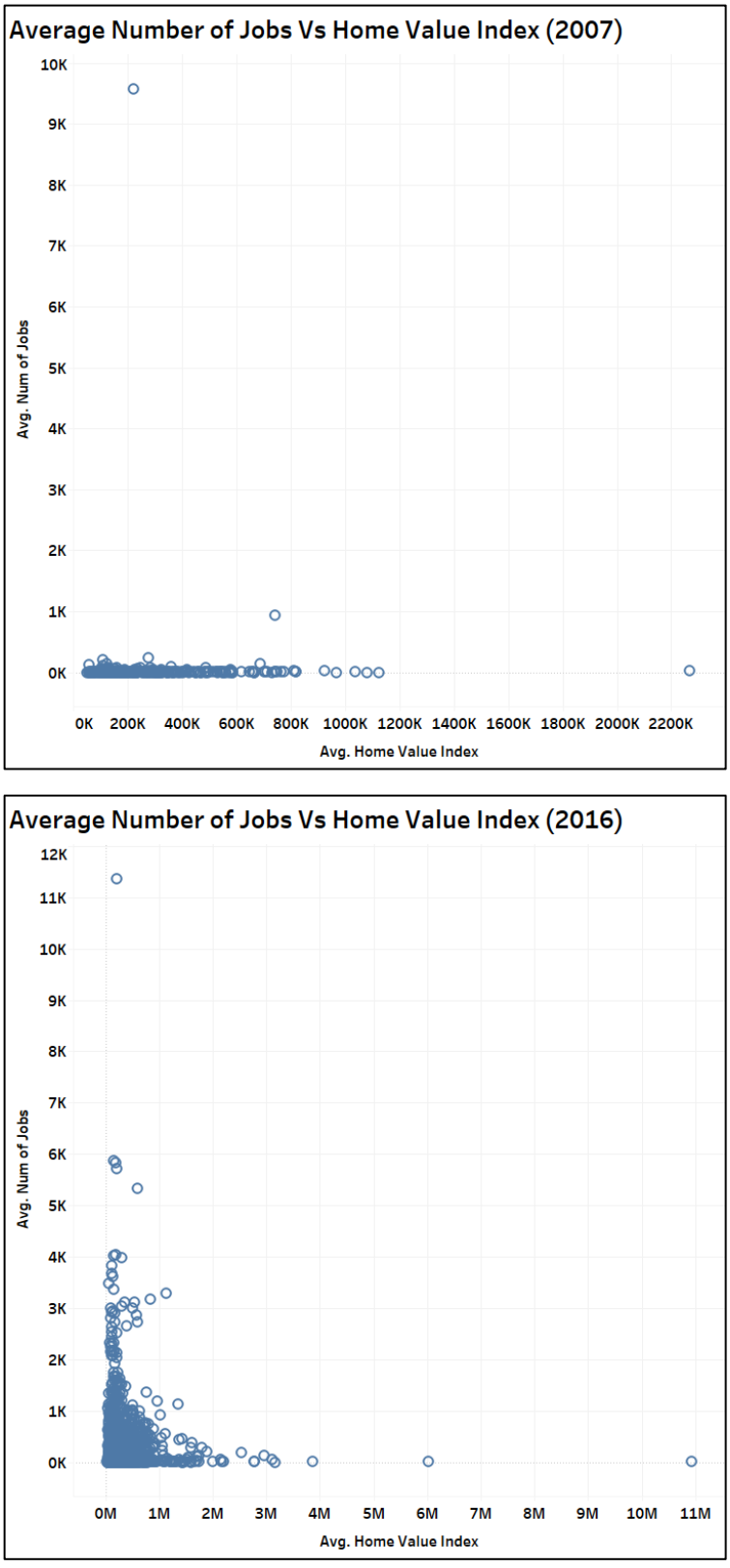

Figure 6: Scatterplot of Average Number of Job postings vs Home Value Index in 2007 and 2016

As we can see above, there is no clear relationship between mean housing price and mean number of job postings (2007 and 2016), though all of the outliers in this category tend to be smaller cities---this is probably just because we dropped some of the major cities because there was incomplete data.

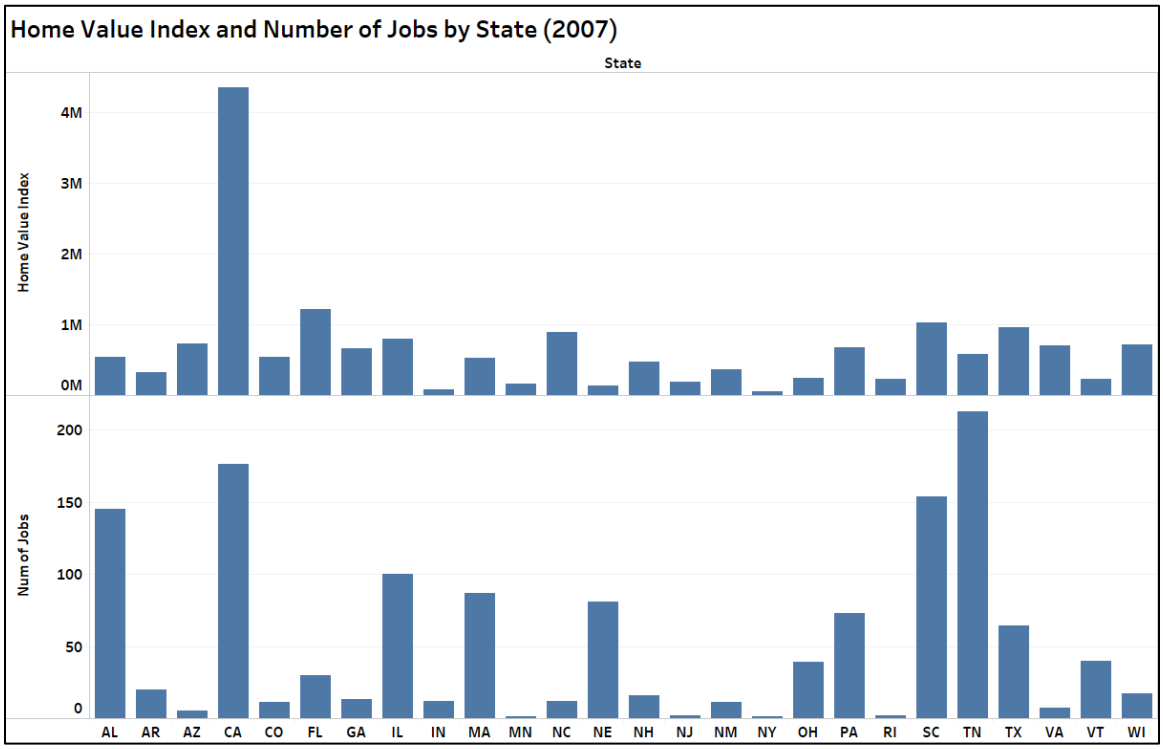

Figure 7: Home Value in Index by State, 2007

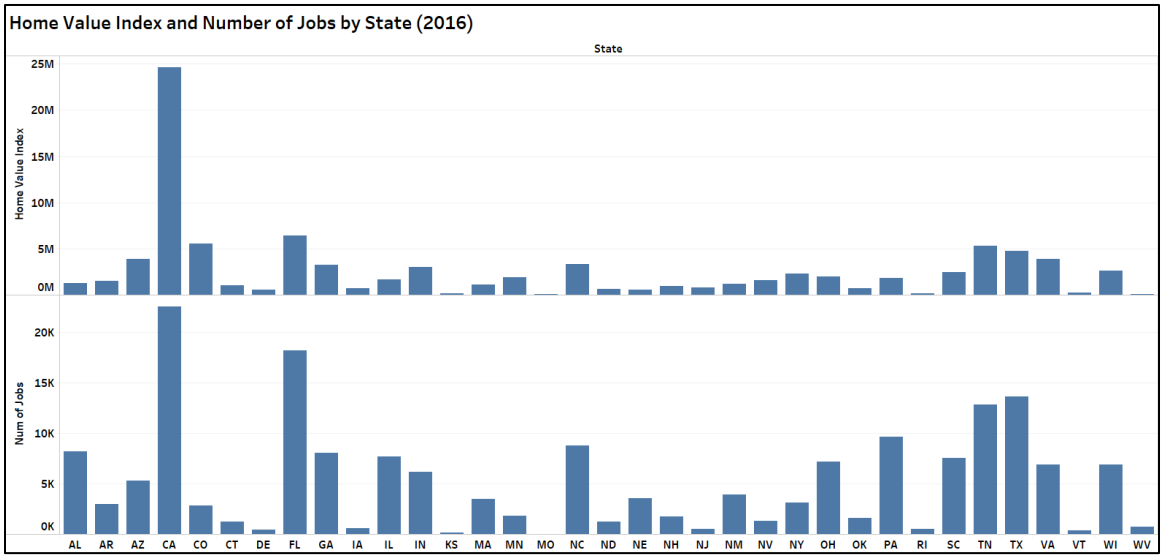

Figure 8: Home Value in Index by State, 2016

In Figures 7-8, we see that the west coast housing market bounced back after the housing crisis, staying far above the over 49 states in this category.