London Housing Case Study

Surya Iyer

Introduction

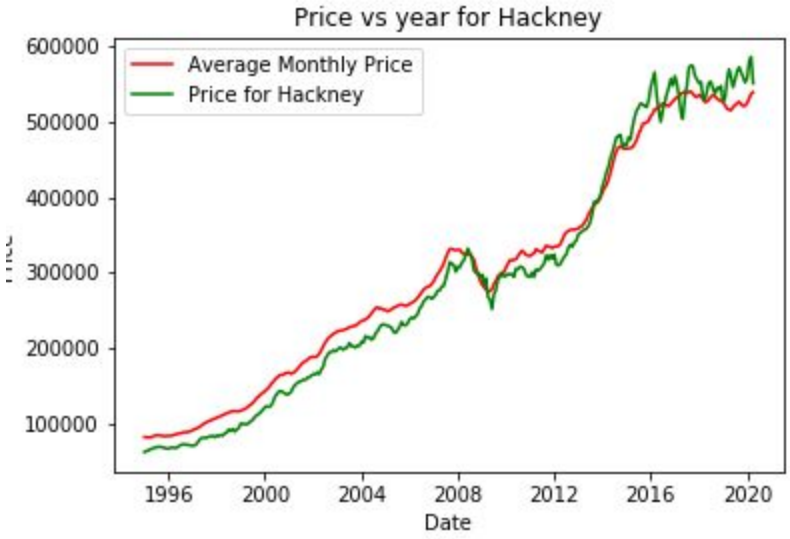
Given a dataset containing prices of houses from 1995 to 2020, we would like to analyze which boroughs have grown in the last 2 decades.

Observations

The overall ratios of prices of the highest and lowest ratios for the last 2 decades is shown in the table below. The overall average increase is about 4.5 times.

Borough	**Ratio of 2018 to 1998 Prices**
Hackney	6.198285561
Waltham Forest	5.834755809
Southwark	5.516485302
Hillingdon	4.20027308
Hammersmith & Fulham	4.137798102
Sutton	4.118522609
Bromley	4.094784685
Harrow	4.059196433
Richmond upon Thames	4.005161896
Hounslow	3.976409106

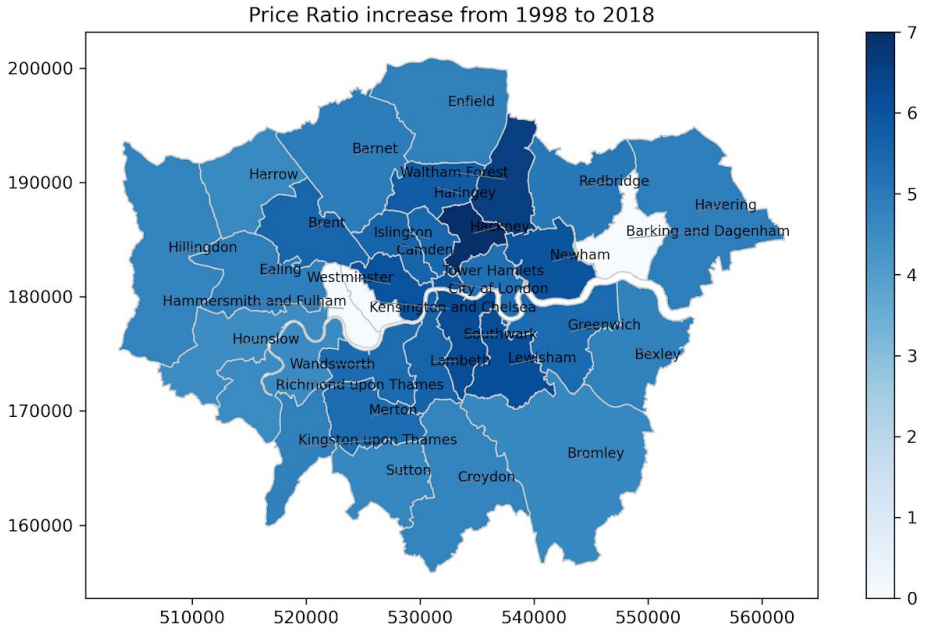
Hackney has the overall largest increase in prices as compared to other boroughs but it is not the most expensive Borough. As we see in the graph, it was below the average monthly prices of houses prior to 2013 and and has grown above the average monthly prices in recent years.



Kensington & Chelsea is the most expensive borough and has grown significantly by 5.08 times since 1998. The corresponding trends can be seen below:-

As we can see the prices of Kensington & Chelsea Borough doubles that of Hackney in spite of it growing the most!

The overall ratios can be seen in the following map: -



Challenges

One of the challenges I faced was to get the ratios in a heatmap. I had to go to external resources and found a really good article of plotting maps into maps using geopandas and matplotlib. The article can be found here https://towardsdatascience.com/lets-make-a-map-using-geopandas-pandas-and-matplotlib-to-make-a-chloropleth-map-dddc31c1983d

Further Investigation:

Looking at the trends, we can see that prices dipped in 2008 which we can account to the recession. I would like to identify why the prices have dipped during in 2018 and why they show a lot of variance in this period.