

An Analysis of London House Prices - INM433 “Visual Analytics”

Edward St John

Abstract—In this research project we aim to explore the different factors behind recent trends in London house prices and sales volume, assessing different factors such as interest rates and Stamp Duty Land Tax reductions (and subsequent reduction extension announcements). The results highlight that these factors do appear to directly influence both price and sales, however caution must be drawn to distorted results during Covid-19 as there are likely to be some effects shown in the market that appear to be due to our factors but may actually be caused by a simple lack of demand and supply due to Covid.

1 PROBLEM STATEMENT

This paper aims to provide an in-depth analysis into the factors affecting UK house prices and sales volume (i.e. the number of houses sold), with a more targeted analysis focused on London specific results. Based on this, we aim to answer the following questions:

- How much do house prices affect sales volume and vice-versa?
- What has been the effect of interest rates on house prices?
 - o Which has more of an impact, decreasing or increasing them?
- How much has the Stamp Duty Land Tax (SDLT) nil rate band increase affecting prices and the number of houses sold, as well as its various extensions?

The data we are analysing is taken directly from the London Datastore which should provide the most reliable average prices and sales volume. The data is also split for us already into all the different boroughs making it very suitable for this analysis, with only some data processing needed to extract this data away from the main table (amongst other pre-processing such as null values) which includes all of the other England regions, making this data very suitable for the questions we are trying to answer. We also combine other data sources (which we'll come onto later in the paper) such as interest rate changes throughout the years, allowing us to answer our second research question.

2 STATE OF THE ART

There have been a few different papers studying the trends behind house prices as well as investigating both internal and external factors behind them. Examples of the former are papers by Hamnett, C [1] and Bracke, P [2] who both look into the differences in price trends for the different boroughs of London. Hamnett, C finds using visual analytics that while the disparity between the top and bottom priced boroughs still stayed intact, the degree of this disparity had lessened due to higher price inflation in the lower priced boroughs. Hamnett puts forward the idea that this is surprisingly due to the increase in demand in more expensive areas, leading to displaced demand in cheaper boroughs resulting in a higher proportional increase in those cheaper areas compared to the higher priced boroughs. Bracke, P looks at this trend from another angle through rent-price ratio calculations. Bracke finds that the rent-price ratios (i.e. the ratio between the median home prices and

annual rent for each borough, essentially used to estimate if it would be cheaper to rent or own property) tend to be lower in more centralised areas, e.g. the more expensive boroughs such as Kensington & Chelsea.

Thinking outside the box, both Holly, S [3] and Shiller, R.J [4] look at outside factors to explain the different trends in house prices. Holly shows that London is heavily affected by other economies mostly due to them being financial centres. In particular, New York has an especially large impact with developments in house prices there playing a role in shocks in London prices. While not dismissing internal and external factors as reasons behind price trends, Shiller, R.J argues that a better reason for the increasing price is in fact a psychological one, claiming buyers perceive rising prices all over the globe which gives them expectations that any property they buy must surely do the same. Shiller also states that they often have the false impression of their house being unique that will become extremely valuable in the future (possibly even more so than the average future price following trends).

This paper aims to build on these 4 pieces of research, looking into other possible factors that could affect house prices that were not researched, such as the effects from interest rates and SDLT within London. We also look into effects on sales volume of houses rather than just the price which should provide a more well-rounded analysis given it seems these papers assume average prices also reflect the added demand for houses and therefore increased volume, which may not always be the case.

3 PROPERTIES OF THE DATA

The house price data is a collection from the London Datastore, released by the Land Registry. The dataset has been processed to split out the different aspects of house prices, separating London borough data into a separate file for analysis given it was dispersed with other England regions. Sales volume has also been recorded, again processed as above to be split.

The London segment of the data dataset contains 10,593 average price values split by month for every year from 1995 to September 2021 and also split by all 33 London boroughs. This is also repeated for average sales volume allowing easy comparisons between price and volume. The rest of the dataset is split by region (e.g. North West, East Midlands etc.) for both average price and number of houses sold – this would be useful for broader comparisons between London prices (and volume) compared to the rest of England to put analysis in perspective.

The data relating to London boroughs was split out to separate files for easier analysis (and linking in our visualisation software). The data was then cleaned to remove null values (the value from the month before was put in their places to avoid assuming any trends before performing analysis) and checked for any missing values / gaps of which there were none.

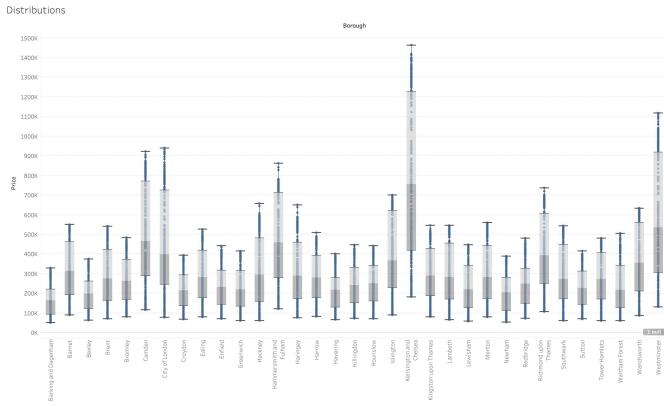


Figure 1 - Data Distribution

It is important to first investigate the properties of our data to assess for the existence of any outliers or peculiar distributions. The boxplots displayed in Figure 1 (a larger version is available in the appendix) allows us to perform this analysis. From first impressions, a few boroughs need to be investigated further with Hammersmith & Fulham and Kensington & Chelsea having a couple of values at the highest end of their respective distributions further away from the other datapoints indicating possible outliers, however after further inspection it was found these values keep in line somewhat with the trends for those specific boroughs. For the most part, the majority seem to have relatively similar distributions with all of the larger distributions coming from the pricier areas – this is to be expected, with the expensive areas more likely to experience a disproportionate amount of variance with abnormally high house prices sometimes distorting the average figures.

A spatial dataset was added from ONS Geography¹ containing the geographical boundaries for each of the London boroughs, being the most accurate one available. This was linked in with our existing dataset to allow a spatial analysis of data distribution in Figure 2. These boundaries were then checked to multiple sources to assess for non-uniformity or lack of coverage, but it was found that the ONS boundary data was accurate and covered all boroughs.

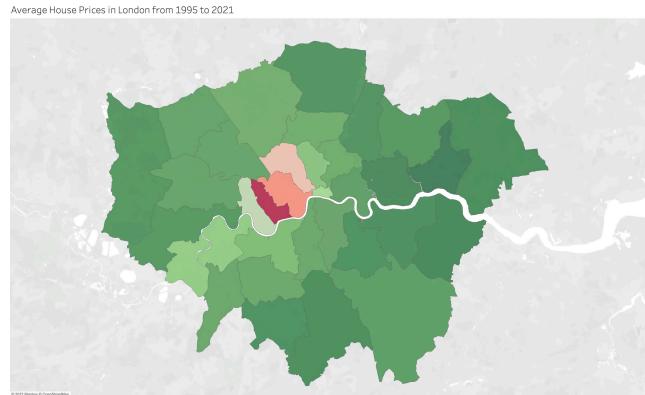


Figure 2 - Average House Prices

Above, we can clearly confirm what we inferred from our boxplots in Kensington & Chelsea and Westminster (the two innermost and reddest boroughs). This large difference in average prices should be taken into considerations when drawing conclusions about London as a whole but may not necessarily change the analysis results.

4 ANALYSIS

4.1 Approach

Before the full analysis of house prices, we first need to set out our methods and approach towards the analysis, showing how a combination of visual analytics and human reasoning paired together can help us answer our analytical questions. The first stage of our analytical approach will be a temporal distribution of house prices across all the different London boroughs in order to provide proper context for the analysis, allowing an immediate visual answer as to where the highest price houses are grouped. Human reasoning will then be used to assess how best to approach the next stage of the analysis and provide the most effective visual analytics going forward – this should be due to the fact that having an initial idea of where the data and results may be centred around (and also potential affects from certain boroughs skewing the overall data in a certain direction) will allow us to alter our analysis results should any particular distributions skew the results one way or another too much.

The next stage of our analysis is to try and assess whether the change in house prices and volume over the years can be attributed to specific external factors (as opposed to general house price inflation). New data for these factors will need to be fed in and attached to our current data before performing the visualisations and subsequent analysis. Firstly, will be looking at the effect of interest rates, using a temporal analysis following the rate changes so assess any correlations with prices. Following this, we will investigate the effects of another factor outside of the housing market using the same visualisation process with Stamp Duty Land Tax (SDLT), or rather the reduction of SDLT during lockdown in the UK due to Covid-19 [5]. We will also analyse the various extensions to the deadline for SDLT reverting back to its previous rates, displaying adjacent temporal graphs for both SDLT and house prices where human reasoning can be used to assess whether any price changes could be due to the extensions. Human

¹ Contains National Statistics data © Crown copyright and database right 2012,
Contains Ordnance Survey data © Crown copyright and database right 2012

reasoning will again need to be used to assess if any observed correlations were actually due to the target variable of rate changes or whether it could be due to different external factors such as the ones put forward by Holly, S and Shiller, R.J.

These results will then be collected and presented together to provide a collective well-rounded view of the factors behind these recent trends in both house prices and the number of houses sold.

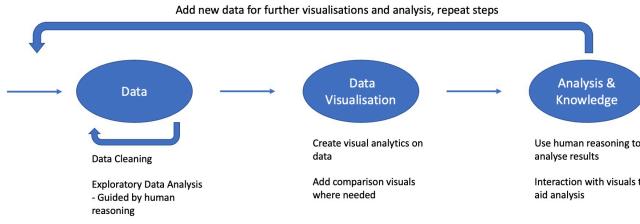


Figure 3 - The Analysis Process Visualised

4.2 Process

For this section, we will be performing the tasks laid out in Section 4.1 to answer our analytical questions. The stopping condition will be when we are able to definitively answer these questions, as well as conclude on the various factors behind the trends that we observe (i.e. if interest rates and Stamp Duty do in fact play a part in prices and sales).

4.2.1 Price vs Volume Trends

Firstly, we will be looking into how much prices affect demand for houses, and subsequently the differences in elasticity for each of the different boroughs to see if any areas are significantly more / less reactive to prices than others. So firstly, let's take a look at London as a whole to see if there is any general correlation before delving into borough-specific correlations.



Figure 4 - London Average Price vs Houses

As we can see in Figure 4, there seems to be a correlation in 2008 / 2009 where the average house price dips as the number

of houses sold ('Volume') goes down although this is most likely due to the 2008 crash [7], however for the most part there doesn't seem to be much of a relationship there. Based on this, it seems a good idea to dig deeper into other factors affecting both of these variables and it doesn't seem we can see many trends from comparing these two in isolation with each other.

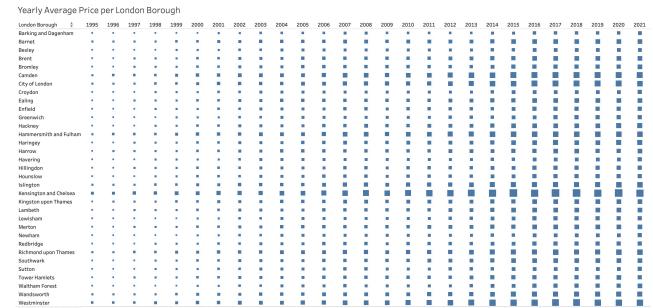


Figure 5 - Trend of House Prices for London Boroughs

We are using the above Figure to gain an initial idea of the price trends for each borough (a larger visual is available in the Appendix). While some boroughs stick out with large price representations such as Westminster, however some boroughs go under the radar such as Wandsworth and Ealing providing relatively large increases for the time period.

4.2.2 Interest Rates

As for factors that could affect house prices, we are first going to investigate the effects of interest rates. Interest rate data was obtained from the Bank of England [6], which we analysed by splitting up two graphs into looking at increased interest rates (Figure 5) and decreased interest rates (Figure 6). We are comparing house prices rather than the number sold as interest rates should affect the price of the house people are looking to purchase more, rather than the actual number of people purchasing.



Figure 6 - The Effect of High Interest Rates

In the Figure above, red lines have been added to indicate points where interest rates have been increased to their highest

points in recent times. Although the overwhelming trend of house prices in London can be seen to be steadily increasing, we are going to investigate a little deeper to spot smaller effects of the interest rate changes – because of this short-term nature of results from high interest rates in this example, human reasoning is needed to assess any correlation. Looking into the red lines above, we can spot pullbacks (indicated above with arrows) from the average prices following rate increases. Interest rates directly affect mortgage rates (being the main factor other than credit history, income etc) offered to prospective buyers, and a larger interest rate for buyers means that they will have to pay more in order to borrow the money needed to purchase the same property, so they will usually aim to buy cheaper properties to compensate for this extra cost. A higher rate may also make current homeowners sell their properties, as if they are on an adjustable-rate mortgage then they too will have to pay this new increased rate. Both of these factors lead to a relatively quick effect on the housing market of slowing down demand (with a small increase in supply) and pulling down the average price of houses or at least stopping their growth, as we can see with each of the points above. The fourth red line above around 2008 shouldn't be fully relied upon to be a direct influence of interest rates as that is around the time of the economic crash of 2008 [7], so it's unclear how much of an effect the high interest rate had there especially since it was reverted back to low levels so quickly after. This is when human reasoning is needed to assess if any perceived correlation is definitely caused by the target variable.



Figure 7 - The Effect of Low Interest Rates

Turning our attention to low interest rates, the effect is a lot clearer to assess compared to high interest rates. Looking at Figure 7, there is a clear burst where prices increase following the lowering of interest rates on all four occasions (a much larger impact than high interest rates, something we wanted to find out in our research question in Section 1). This is due to it being cheaper for buyers to borrow money, and so will borrow more money to purchase a more expensive house than they would if the interest rate was higher. This leads to more expensive houses having higher demand, pushing the average price of properties up since there is more demand for the same amount of supply. As you may have spotted, there is not an

immediate effect on the house prices unlike the effect of when interest rates are raised as in Figure 5. This lag in showing the effect on the market is due to people taking time to apply for mortgages following the rate changes, and so delaying the impact on house prices for a few weeks (displayed by the difference in where rates were dropped at the green lines and where prices actually start to increase again). This again highlights the need for human reasoning in analysis rather than simply use computational methods, a combined approach like this allows more reliable analysis.

4.2.3 Stamp Duty Land Tax

Seeing as interest rates appear to have a lasting effect on house prices, other potential external factors should be considered that might also affect house prices and the number of houses sold. Stamp Duty Land Tax (SDLT) [5] is one in recent times that might have had a large effect. SDLT is due to be paid on properties bought over a certain threshold in England and Northern Ireland. However, due to Covid-19 and the trying to stimulate growth as a result of potential economic downturn, the nil rate band for SDLT (i.e. the value for a property at which there is no SDLT) was increased in July 2020 until the end of March 2021. In March 2021, they then extended the deadline for SDLT to the end of June 2021. This reduction means property buyers can now buy pricier houses within the nil rate band paying no tax, as well as houses above this nil rate band being cheaper overall (as the tax isn't applied as a blanket for the whole value, it's done at the value above the nil rate band).



Figure 8 - Effects of Stamp Duty Land Tax

In Figure 8 above, we investigate the effects of SDLT by assessing changes to both house prices and number of houses sold at three timeframes:

1. The start of the SDLT nil rate band increase (green line)
2. The announcement for an extension of the increase (yellow line)
3. The end of the SDLT extension (red line)

As soon as the SDLT exemption is announced, there is a clear uptick in both sales and price. With the extra savings from

less tax being paid, it's clear people can afford higher priced houses than they would have before, pushing up the price (there was nearly a 2% increase in average house price in the following month alone). Also, people who were thinking of buying a house within the next year or two would have now pushed that forward to take advantage of this tax break which increases the number of houses sold as above. Towards the end of the (initial) deadline, there is an even bigger increase in the number of houses sold in the final month presumably due to trying to get house sales over the line and completed before the old SDLT laws come back into effect, resulting in over a 58% increase in sales in March alone. Once the extension was announced, there was an initial drop (which is to be expected, as most of the houses that would have been sold in the period after the deadline would have been moved forward to take advantage of the tax relief), followed by a similar pattern with an even larger relative increase up to the final deadline date of June 2021 (with nearly double more houses being sold than the penultimate month, from 5,639 to 10,522 houses). Finally, as expected, the number of houses sold drops off hugely after the final deadline. Unfortunately our sales volume data doesn't cover past July 2021 so no analysis could be performed on how well the market recovered after the SDLT laws were reverted back. It's clear from the above figure that these changes seem to be stemming from these policy changes, rather than any other specific external factors.

4.3 Results

Following our analysis, we have managed to answer our research questions set out in Section 1. For the first question, there doesn't seem to be too much of a correlation between house prices and how many are sold (with the exception of the Financial Crash of 2008 showing both decreasing significantly). With our other question regarding interest rates, we have shown correlations both with high interest rates restricting growth through lower purchasing power for buyers, and lower interest rates having the opposite effect with buyers able to borrow more money (with decreasing interest rates having a more substantial intended impact). The implications found by this should help policymakers estimate the effects of interest rates in the housing market, and also should help prospective house buyers in assessing when the right time is to buy houses. In terms of the effect of SDLT, its clear from showing the timings of announcements that it has clearly had a fairly large effect on the market, especially as we near the deadline(s) with a flurry of sales. However, these events due to SDLT can distort the trends in the short run, so should this analysis be performed again this should be considered.

5 CRITICAL REFLECTION

The use of both interest rates and SDLT rates provided a whole other angle to our analysis in this paper rather than just relying on trends between house prices and number of houses sold in different boroughs. Unfortunately, there were too many boroughs to be able to present both house price and number of houses sold on the same graph which would have aided visual analysis. However, a reasonable amount of visual analysis was still managed to be undertaken through the line charts as shown in most of the figures above. During this paper there was more of a focus on the factors affecting house prices and volume sold with a smaller amount of analysis performed on price and volume data in isolation – more could have been done on this

topic however with length constraints focus had to be applied and the external factors proved to be the more interesting of the research questions to answer in fuller detail.

The data we used had several limitations, one of which is when using it in isolation past 2019 makes it hard to see which changes are natural trends in the market and which were caused by the effects of Covid-19 (which is itself part of the trend of course but does distort the natural trend set by previous years). In terms of data availability, there was also a mismatch between the amount of data provided for house prices compared to number of houses sold, with house price data available up to September 2021 whereas house volume data only available up to the end of June 2021 (explaining why our SDLT analysis had to be cut short with a lack of data past July). Past July, there was still a smaller nil rate band exemption still in place between July and September 2021 which would be interesting to analyse in future research. For future papers doing the same line of analysis this shouldn't be an issue, as hopefully London Datastore will have updated their House Price data to include these months to allow that analysis to happen. Also, for both the interest rates and SDLT analysis, all the boroughs were aggregated which may not have been the best representation of London as a whole if there was an abnormal trend for one of the boroughs which disrupts the average. A more detailed analysis would be able to inspect each brough individually to see the different effects of these factors, however this paper focuses on covering multiple factors. While these limitations did hinder the analysis, answers to the research questions set out in Section 1 still had a reasonable amount of evidence behind them providing valuable insights.

Table of word counts

Problem statement	222
State of the art	432
Properties of the data	500
Analysis: Approach	417
Analysis: Process	1497
Analysis: Results	198
Critical reflection	444

REFERENCES

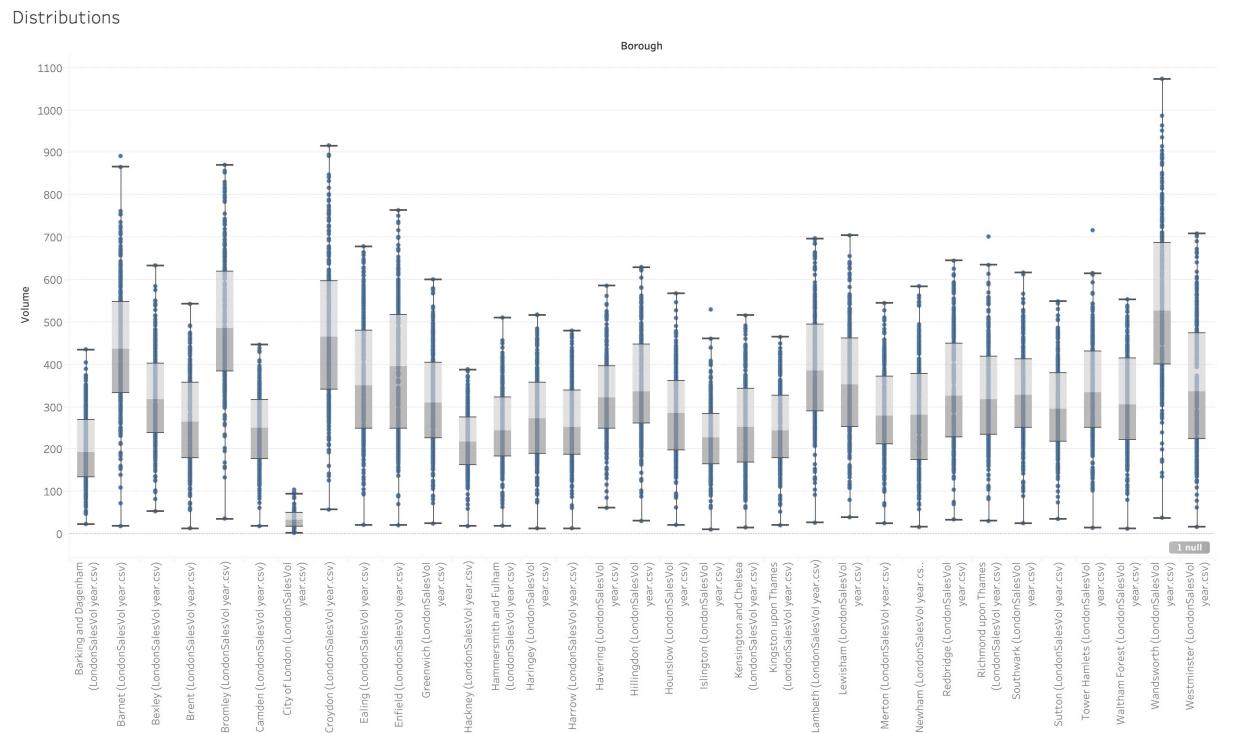
The list below provides examples of formatting references.

- [1] Hamnett, C., 2009. Spatially displaced demand and the changing geography of house prices in London, 1995–2006. *Housing Studies*, 24(3), pp.301-320.
- [2] Bracke, P., 2015. House prices and rents: Microevidence from a matched data set in Central London. *Real Estate Economics*, 43(2), pp.403-431.
- [3] Holly, S., Pesaran, M.H. and Yamagata, T., 2011. The spatial and temporal diffusion of house prices in the UK. *Journal of urban economics*, 69(1), pp.2-23.
- [4] Shiller, R.J., 2007. Understanding recent trends in house prices and home ownership.

- [5] Gov.uk., 2020. Stamp Duty Land Tax: Temporary Reduced Rates. Available at: <https://www.gov.uk/guidance/stamp-duty-land-tax-temporary-reduced-rates> (Accessed: 16/01/2022)
 - [6] Bank of England., 2021. Interest Rates and Bank Rate. Available at: <https://www.bankofengland.co.uk/monetary-policy/the-interest-rate-bank-rate> (Accessed: 16/01/2022)
 - [7] Wu, D.D. and Olson, D.L., 2015. The Real Estate Crash of 2008. In *Enterprise Risk Management in Finance* (pp. 23-31). Palgrave Macmillan, London.

APPENDIX

Larger visualisation of data distributions:



Larger visualisation of price trend per London borough:

Yearly Average Price per London Borough