Understanding the Spatial and Temporal Impact of the Coronavirus 2019 Disease (COVID-19) on House Prices in England and Wales

William Dee

Abstract— This study investigates how annual house price appreciation has been affected during 2020 by the COVID-19 pandemic. It uses a dataset-fusion approach, combining price-paid data with COVID-19 cases, population and average garden size figures, aggregated at a county level. Using visualizations to inform the process, new features are extracted to be modelled using regression techniques. A geographically weighted regression (GWR) model was found to be most accurate (R^2 of 0.73), confirming the importance of the spatial attributes of the data. The results demonstrated clear temporal trends, identifying stark periods of low appreciation during lockdown, and a recovery after the August Government stimulus measures. The trends were inconsistent with prior periods and correlate with the additional datasets added. Spatially, clustering resulted in four distinct groups of counties that experienced the pandemic differently in terms of house price appreciation, giving insight into the preference changes of buyers and identifying counties in need of further support.

1 Problem Statement

Real house prices and their appreciation over time can be an important economic indicator as well as the driving force behind supply and demand dynamics in the housing market itself[1]. The Coronavirus 2019 Disease (COVID-19) has presented a unique shock to this market in the form of temporally delineated periods of national lockdown, as well as spatially heterogenous case numbers and local lockdowns.

This report will focus on answering the following questions:

1. How has the COVID-19 pandemic affected annual real house price appreciation and the number of houses sold in 2020 compared to prior years?
2. Are there distinct temporal and/or spatial trends in the housing market that can be observed and linked to COVID-19's impact?
3. How accurately can a regression model be fit with publicly-available housing data to model annual changes in house price appreciation?

It is an important study to undertake given the ongoing nature of the pandemic. The effects are of clear importance to the government who were quick to intervene with stimulus[2], in order to avoid a housing crisis, as well as to individual buyers and sellers.

The dataset has been constructed using the UK government price paid data and has been combined with other datasets from the Office for National Statistics (ONS) to obtain date and location-specific information about population, COVID-19 cases and outdoor space size. The resulting dataset is suitable for analysing trends on a district or regional level, but not at a point-specific scale as structural information (i.e., bedrooms, bathrooms, individual garden size) is not included.

2 State of the Art

I utilized two main papers that dealt with similar data. The first, by Holly et al.[4], looked at quarterly UK regional house price changes between 1973-2008. The paper investigated how systemic shocks impact house price changes and how these effects propagate from one region to another. A linear price diffusion model was constructed, alongside spatio-temporal response functions, to determine to what extent a dominant region (found to be London in the UK) can spread the impact of a shock.

Shocks were shown to decay more slowly over space than time, as Holly et al. found that London could recover from a shock within 2 years, whilst surrounding regions took longer to recover the further they were from London.

The paper impacted my approach in several meaningful ways. Firstly, it confirmed that shocks do not impact house price changes evenly across geographical or time dimensions. This gave credence to my research questions and the possibility of constructing a worthwhile model. However, I am making the assumption that COVID-19 is a comparable shock to the ones analysed in the reference paper, which are determined as functions of price changes rather than individually-specified events.

Both this paper and the second paper emphasized the importance of converting nominal house prices to real house prices, to adjust for monetary inflation. Both papers also utilized house price changes, rather than stationary house price values which "are only snapshots of the property values in a specific time window"[5].

The research also guided me towards data aggregation as it showed meaningful conclusions can be reached by looking at house prices on a regional basis (11 defined UK regions).

The second paper, by Kang et al.[5], evaluated how combining different datasets, such as structural house attributes, house and street-level photo's, transportation links, local amenities and socioeconomic data could be used to construct regression models at both the fine-scale point level (individual houses) and the neighbourhood level, to predict house price appreciation. 20,000 houses in the Greater Boston area were included in the analysis.

This paper is more focused on visual analytics than the first and produced heat-maps at the neighbourhood scale to visualize the dataset variables and model results, which I have replicated. The approach also guided me towards using the state-of-the-art Geographically Weighted Regression (GWR)[6] technique, and comparing it against a multiple linear regression (MLR) model. GWR uses coordinate data as an additional input variable and has been found to yield superior efficacy when observations exhibit spatial heterogeneity[7,8], as basic regression models assume observation independence which does not hold for spatially-related data. This follows Tobler's First Law of Geography[9] that "everything is related to everything else, but near things are more related than distant things."

Whilst my dataset does not have the point-scale data that Kang et al. had access to, I can follow a similar process at a UK county-level to their neighbourhood analysis. I am assuming that adding in the COVID-19 case data as well as explicit lockdown periods provides enough spatio-temporal context to my model so that its efficacy is comparable to the reference paper.

3 Properties of the Data

5 different datasets were combined to form the final input used:

HM Land Registry Price Paid Data for England and Wales[10] – contains all individual property sales between 1st January 1995 and 31st October 2020 that were lodged with the HM Land Registry.

It includes 25.5 million datapoints across 14 columns, including the nominal price paid, as well as the categorical columns; date sold, property-type (e.g., flat or detached), new-build or not, and leasehold or freehold. 8 other columns specify varying degrees of spatial locality (e.g., street, town, county) and so the resolution can be user-specified.

UK Government Coronavirus Case Data[11] – COVID-19 case data recorded by specimen date at an Upper Tier Local Authority Level (similar to UK County level). Includes 54,652 datapoints, being the daily recorded cases for each location.

ONS Access to Garden Space[12] (April 2020) – compiled through Ordnance Survey geospatial analysis of address data combined with large-scale topographic data[13]. 8,480 numerical datapoints measure the average private outdoor space for residential properties, summarized at a regional level.

ONS Population Estimates[14] (June 2019) – Regional population estimates, constructed by annually updating prior census numbers (last taken 2011) using GP registry data and surveyed changes to local populations.

ONS Retail Price Index (RPI)[15] – Monthly official government inflation data selected between January 1995 and October 2020.

The datasets were merged based on regional identifiers (e.g., Greater London) and date. Some of the datasets had alternative ways of categorising or spelling regions and so manual intervention was required to ensure all datapoints were allocated similarly to the base HM Land Registry's County-level tiers.

The price paid data was incomplete after October 2020 due to the delays in uploading information and so this was determined as the cut-off date to analyse for this project. There were several large outliers discovered in the original dataset which identified that property type 'Other' referred to commercial properties, and so these values were excluded.

Other outliers were searched online to verify accuracy, the largest of these being a £160 million flat that was actually a refinancing rather than a sale[16] and so was dropped from the data. I have made the assumption that the remaining sales are valid, and that, if there are any other inconsistencies, the collection methodology is still applied consistently and so yearly comparisons retain legitimacy.

UK COVID-19 case data was incorrectly recorded[17] between September 25th and October 2nd. This is why cases at specimen date (i.e., date tested) were chosen for this analysis, rather than at date recorded. Location was initially determined by patients NHS registered address, but the methodology was switched on 16th November to address given at point of testing. Therefore, it must be noted that this report reflects the older recording methodology.

Histograms were plotted for each transformed numerical feature in the combined dataset to understand their distributions (Fig.1).

As was done in the reference paper[5], real price was converted to its natural logarithm which was more normally distributed compared to the skewed raw data (Fig.1). For regression modelling; population, COVID-19 cases and average garden size were log-transformed also.

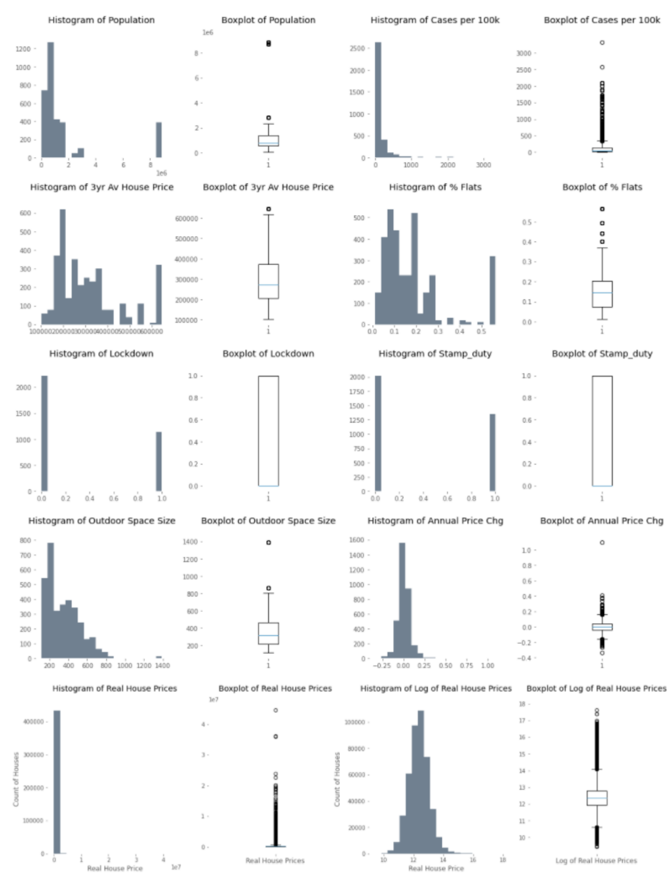


Fig. 1 – Distribution of Transformed Variables

4 Analysis

4.1 Approach

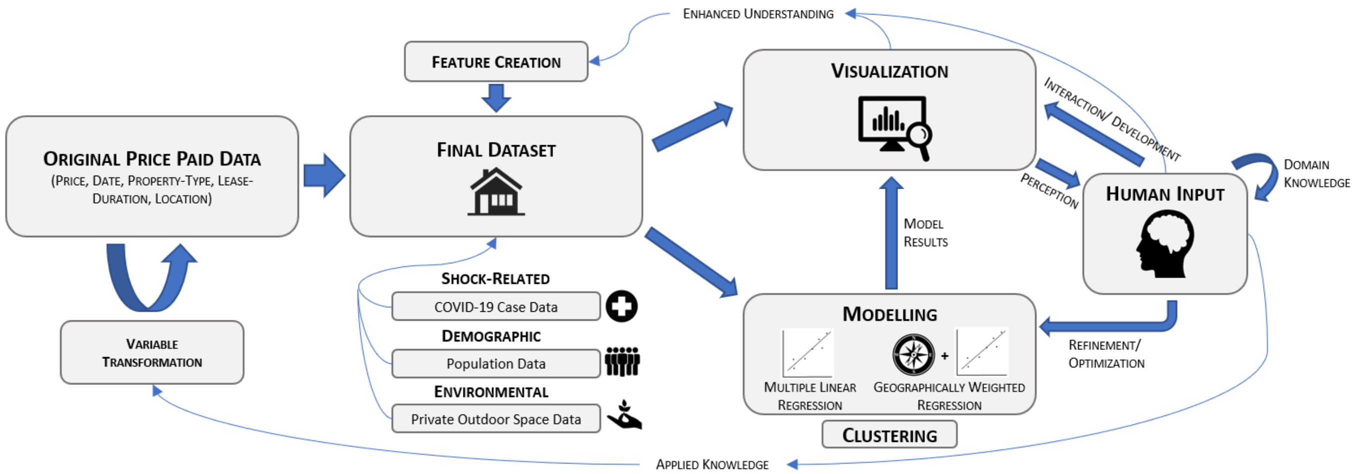


Fig. 2 – Workflow Diagram of this study.

The workflow followed has been outlined schematically in Fig.2. This section describes how visualizations combined with computational methods informed the evolutionary process of this study, creating a feedback loop through human reasoning to develop an understanding of the data. The steps described below are illustrated in Section 4.2 – Process:

A) **Establish if 2020 house price changes exhibit any clear temporal differences to prior years**: After pre-processing steps, high-level trends are analysed using the base housing dataset.
Computational: The price paid column will be transformed from nominal values to real prices, then features created for annual monthly percentage change and number of houses sold.
Visualization: Time series plots will be used to evaluate these differences. Patterns noted will inform the next analysis steps.

B) **Compare London to Ex-London**: Informed by my own domain knowledge, as well as Holly et al.[4], which suggests that London, as the dominant housing region, is likely to exhibit different temporal characteristics than other areas of England and Wales.

C) **Investigate Trends for Different Property Types**: Driven by human reasoning that flats may have become less attractive to buyers during lockdown[18].
Computational: Group house sales by property type and calculate percentage change metrics.
Visualization: Plot time series with a higher resolution (focusing on 2020) to observe temporal shifts in preference throughout the year. This can inform feature creation if strong preferences are observed.

D) **Examine Regional Differences**: Analyse the spatial components of the house price data.
Computational: Group data by counties and month in 2020 and merge with the geo shape file of England and Wales.
Visualization: Plot choropleth maps, using any temporal trends identified in steps A-C as a focus of the visual analytics.

E) **Integrate other Datasets**: The outdoor space, population and COVID-19 case datasets were chosen to complement the house price data as all three have featured in the news as reasons for changes in house prices during 2020[19,20].

F) **New Feature Creation**:
Computational: Create new features based on trends observed so far (i.e., binary 'flat or not' column), and added datasets (i.e., COVID-19 cases per 100,000).
Visualization: Plot a correlation matrix of all dataset features to determine correlation with target variable (annual real price change), and co-linearity with each other. This will inform how appropriate the combined dataset is for regression modelling and clustering.

G) **Regression**:
Computational: Perform MLR and GWR methods and evaluate performance of predicting annual percentage price change with the R^2 metric. Employ human reasoning to adjust the model parameters and refine the inputs. R^2 will be compared to the reference paper and comparable performance can be considered one stopping point.
Visualization: Scatterplots of actual vs predicted values will allow visual interpretation of the model fit whilst

residual plots can detect non-linearity, hetero-skedasticity and outliers in the error terms.

H) **Clustering**:
Computational: Standardize the input features and utilize K-Means clustering to group counties according to the features established in steps A-F.
Visualization: Apply silhouette plots as well as choropleth maps to decide on the appropriate number of clusters. Coherent clusters that inform the research questions will represent a second stopping criteria.

4.2 Process

This section will detail the process following the steps outlined in Section 4.1 Approach:

A) **2020 - Visual Comparison to Prior Years:**

Following pre-processing, the data was aggregated monthly from January 1995 to October 2020. A monthly summary was considered appropriate as this timeframe is quoted in government housing reports[21], and annual monthly change is cited frequently in media property news[22]. Furthermore, RPI figures are released monthly and so this allows for the most accurate conversion from nominal to real price.

Despite the conversion to real price, the average real house price is ~£350,000 today versus ~140,000 in 1995 as house prices have generally increased above inflation. Monthly real prices were converted into annual change percentages to allow for a more explicit comparison between years (Fig.3).

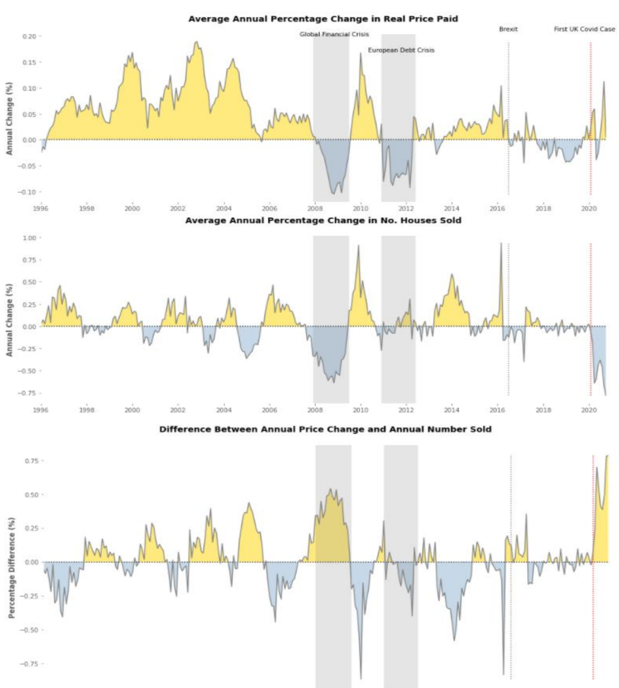


Fig. 3 - Time Series Charts (Jan 1995- Oct 2020)

The plot has been annotated with major events that have been linked to house price changes over time. Unlike the Global Financial Crisis and the European Debt Crisis, annual price changes in 2020 have generally not declined. However, the decline in monthly number of houses sold eclipses both these crises and represents the worst period in the time series. Plotting the difference between these two measures reveals the unusual nature of the 2020 housing market as there is the biggest positive difference observed in 25 years, which clearly requires further explanation.

B) Compare London to Ex-London:

London was compared to all regions excluding London in the same manner as step A. The time series chart (Appendix A) revealed that London annual price changes are generally more volatile and the region recovers more quickly from the shocks labelled in Fig.3, corroborating the findings of Holly et al[4]. This step confirms that geography plays a key role in determining the extent of house price reactions to a shock, and the duration of that impact. The number of houses sold declined similarly in both regions. Given that there appears to be less spatial variation in number of houses sold, and a decline is in line with prior periods of shock according to the time series, the focus of this report will be on the abnormal annual house price changes observed.

C) Property Type Trends:

The dataset was divided into the four different property types – flats, detached, semi-detached and terraced. Time series were plotted for the last 3 years to allow greater resolution for my interpretation. As hypothesized, flats exhibited more negative annual price changes than the other types during 2020, despite similar changes during 2018-2019 (Appendix B). To build on this finding, flats and detached houses were compared during 2020 (Fig.4).

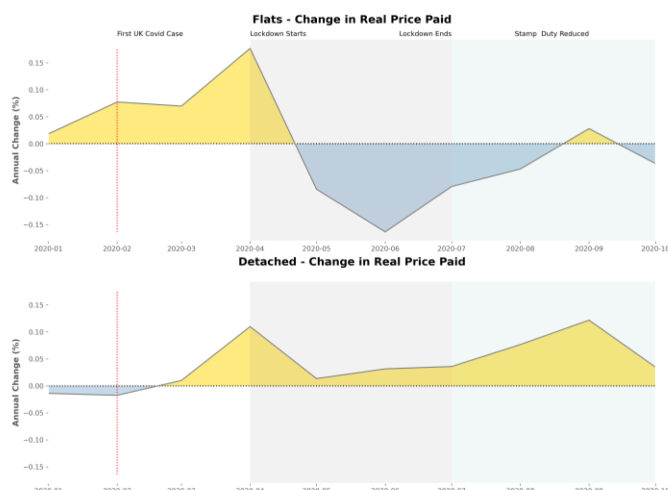


Fig. 4 - Time Series showing percentage annual price change for flats vs. detached properties in 2020

Fig.4 shows clear differences between two periods – being during the National Lockdown (23rd March-4th July) and after SDLT rates were reduced (8th July onwards). During lockdown prices generally decreased from March levels for all types of properties, whilst after July there is a noted improvement in annual price changes.

D) Regional Differences:

To investigate the spatial and temporal differences observed in steps A-C more deeply, the dataset was aggregated at a county level and changes in annual price were plotted on choropleth maps (Fig.5). Given London exhibited differences to the rest of the regions in Section B, this is an important step to understand if there may be clusters of regions that experienced similar price changes to one another.

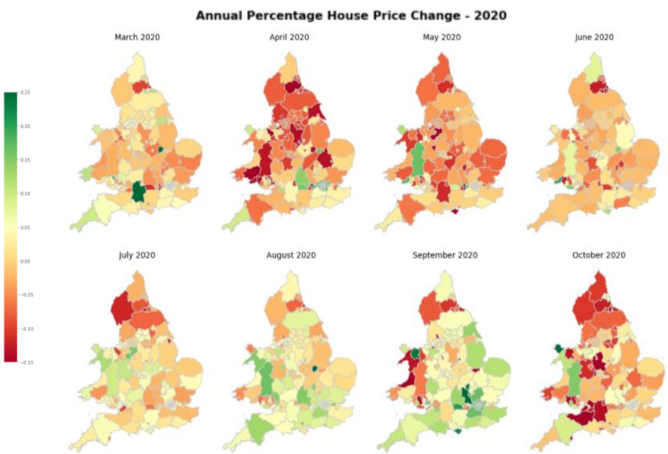


Fig. 5 – Choropleth maps showing percentage annual price change according to county and month during 2020.

This step follows the guiding principles in "Strategies for the Visualization of Geographic Time-Series Data"[23], that multiple views, exhibiting both time, as well as space, scaling, are integral to interpreting spatio-temporal data.

The charts above confirm that during lockdown (March-June) most counties annual house prices fell, whilst after July they began to recover, likely aided by the SDLT reduction. October counters this trend, suggesting that a more granular measure of COVID-19 cases (see Section E) may prove additive to our understanding as I know cases increased dramatically that month.

The charts demonstrate some spatial association exists between the price changes in different counties. There is a divide between the North-West and the South-East as the former was more negatively affected during the lockdown period and the latter rebounded more strongly after SDLT was reduced. Clustering may prove insightful to understand this dynamic further, and regression based on well-chosen spatio-temporal features could yield positive results.

E) Integrate other Datasets:

Steps A-D have highlighted differences between 2020 and prior years, as well as differences in regions, months of the year and property types that appear to align with COVID-19 impact periods. To investigate this further the Government COVID-19 cases dataset was merged with the existing dataset by month and county.

To investigate news article claims of an "exodus" from busy cities[19], and an increased preference for outdoor space[20] caused by the pandemic, the ONS Population and Private Green Space (i.e., gardens), datasets were also merged.

F) Feature Creation:

The COVID-19 case data contains raw numbers of positive tests per county. This has been transformed into cases per 100,000 population to be in line with government and news reporting statistics, and to provide a more accurate comparison across counties. Fig.6 provides a comparison to Fig.5 in visually linking COVID-19 cases with annual house price changes.

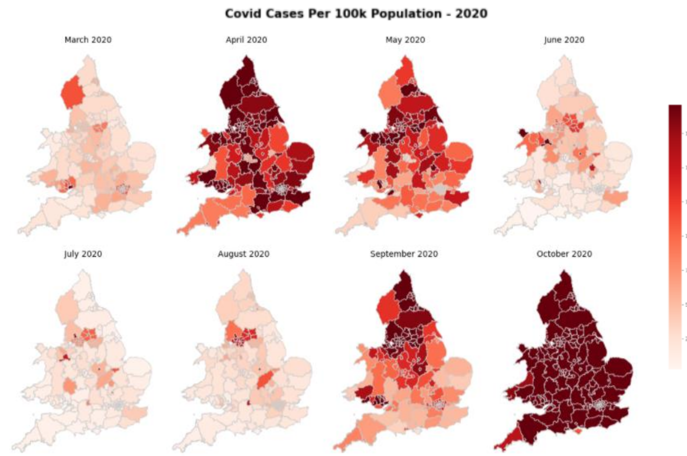


Fig. 6 – Choropleth maps showing COVID-19 cases per 100k population according to county and month during 2020.

Case numbers were generally higher when house prices were declining year-on-year. The high case figures in October may explain why house prices again began to decline, despite the lower SDLT. The differences in cases numbers across the country exhibits a similar North-West to South-East divide that was observed in Step D, with some groupings in key Northern and Welsh cities.

The distribution of house prices in 2020 was compared to the 2017-2019 period (Appendix C). Small differences were noted, mainly in September where houses valued approximately £500,000 (the value which yielded the maximum SDLT saving) were sold more frequently. This led to the inclusion of the 3-year average house price per county.

Step C resulted in the creation of a variable which reflected the percentage of flats as a proportion of all houses sold in each county as an average of the last 3 years. The 3-year period from 2017-2019 was chosen because 2016 could be impacted by the results of Brexit[26].

Binary columns representing whether the month was during a national or local lockdown on a county-basis, and whether the SDLT tax had been reduced were also created. The attribution of lockdowns on a county-basis was manually implemented, informed by news reports[25], as an official dataset does not currently exist.

The features were plotted on a correlation matrix (Appendix D), along with the target variable, to assess collinearity and appropriateness to use for regression modelling and clustering.

G) Regression

The regression results confirmed that spatial attributes of the data were an important variable as GWR outperformed MLR, resulting in an R^2 of 0.73 versus 0.64 (Fig.7). The local R^2 map demonstrates that GWR finds the border of England and Wales the toughest region to model. This could be because both countries followed differing approaches to local lockdowns, causing uncertainty about house purchasing at the borders.

The GWR model underperformed the model in the literature[5] (which achieved an R^2 of 0.77), however the underperformance was marginal considering the vastly less granular dataset employed, which had no structural house attributes and street-level photographs. This suggests that, due to the pandemic, annual house price changes during 2020 have become more dependent on their macroscopic spatio-temporal attributes.

Residual plots (Fig.7) were analysed to observe the consistency of each regression's predictions, which were determined to be homoscedastic for both models.

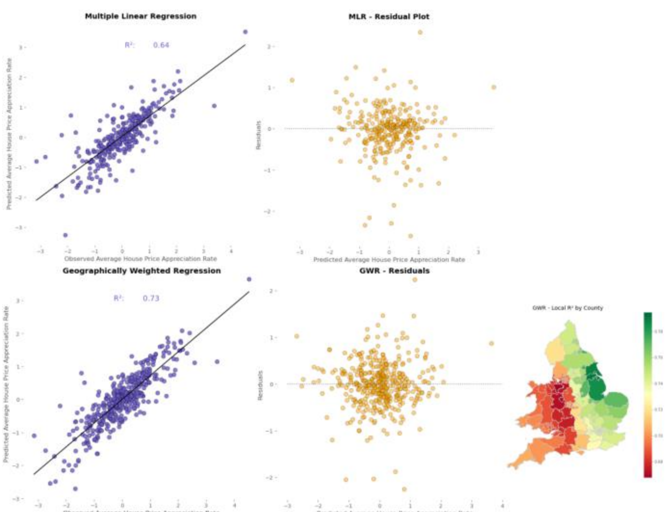


Fig. 7 – Regression Scatterplots and residual plots for MLR and GWR, R^2 shown by county for GWR

The GWR model provides further insight into how geography alters the predictive-power of each variable, allowing feature correlation with annual price change to be visualized regionally (Fig.8).

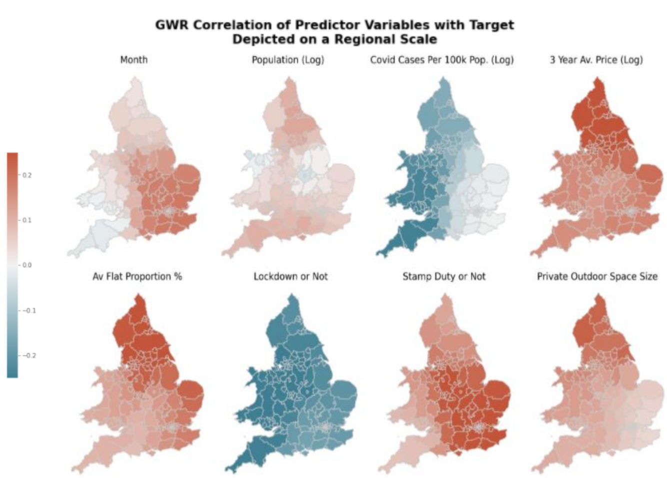


Fig. 8 – GWR Regional Correlation of Predictor Variables with Target

F) Clustering:

The standardized features were transformed using Principal Component Analysis (PCA) prior to clustering, due to multicollinearity (Appendix D). Silhouette scores were calculated and visualized (Fig.9) to determine possible values of k to use for K-Means clustering. Choropleth maps and scatterplots were also used to assess cluster coherence.

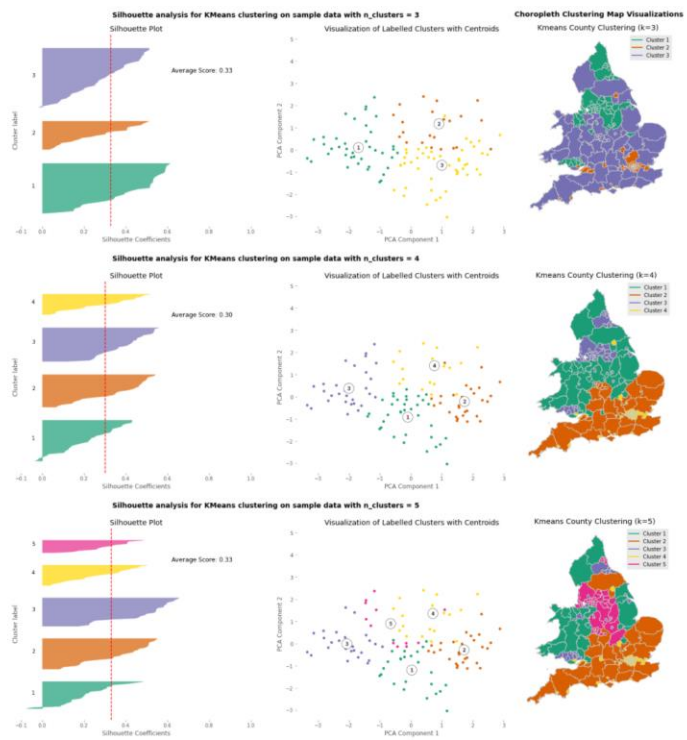


Fig. 9 – Cluster visualization through silhouette scores, scatterplots and choropleth maps

Human judgement is required to determine 'k' as the silhouette scores are roughly equal. Looking at the cluster centroids and the maps I decided that 4 clusters represented the most coherent clustering of the data based on my prior analysis and knowledge.

When evaluating the average variable metrics associated with each cluster, in combination with the maps in Fig.9, it's clear that clusters 1 and 2 represent the more rural areas of the UK, exhibiting the largest average outdoor private space. Cluster 2 also includes the higher-population commuter zones surrounding London and is the highest-priced cluster. Cluster 3 signifies the worst-hit COVID-19 case areas outside London, mostly inner-city areas with small outdoor space. Cluster 4 is largely Greater London, as well as some coastal cities, and has the highest flat proportion. See Appendix E for a full summary of metrics and counties by cluster.

In the 2020 time series plot (Fig.10) there are clear temporal differences observed between the clusters. Cluster 3 has performed the worst, showing the negative impact that higher COVID-19 cases has on annual price changes, especially in the North-West (also evidenced in Fig.8). Cluster 2 was least impacted during lockdown, and rebounded the strongest when SDLT was reduced, further confirming the preference for Southern England and large garden sizes. Cluster 4 also recovered quickly, demonstrating London's ability to bounce-back from a shock. The South-East rebound post-stimulus is also validated by Fig.8.

The comparison to 2019 shows the exceptional volatility of price changes during 2020, as well as stark differences in temporal trends behind these changes.

4.3 Results

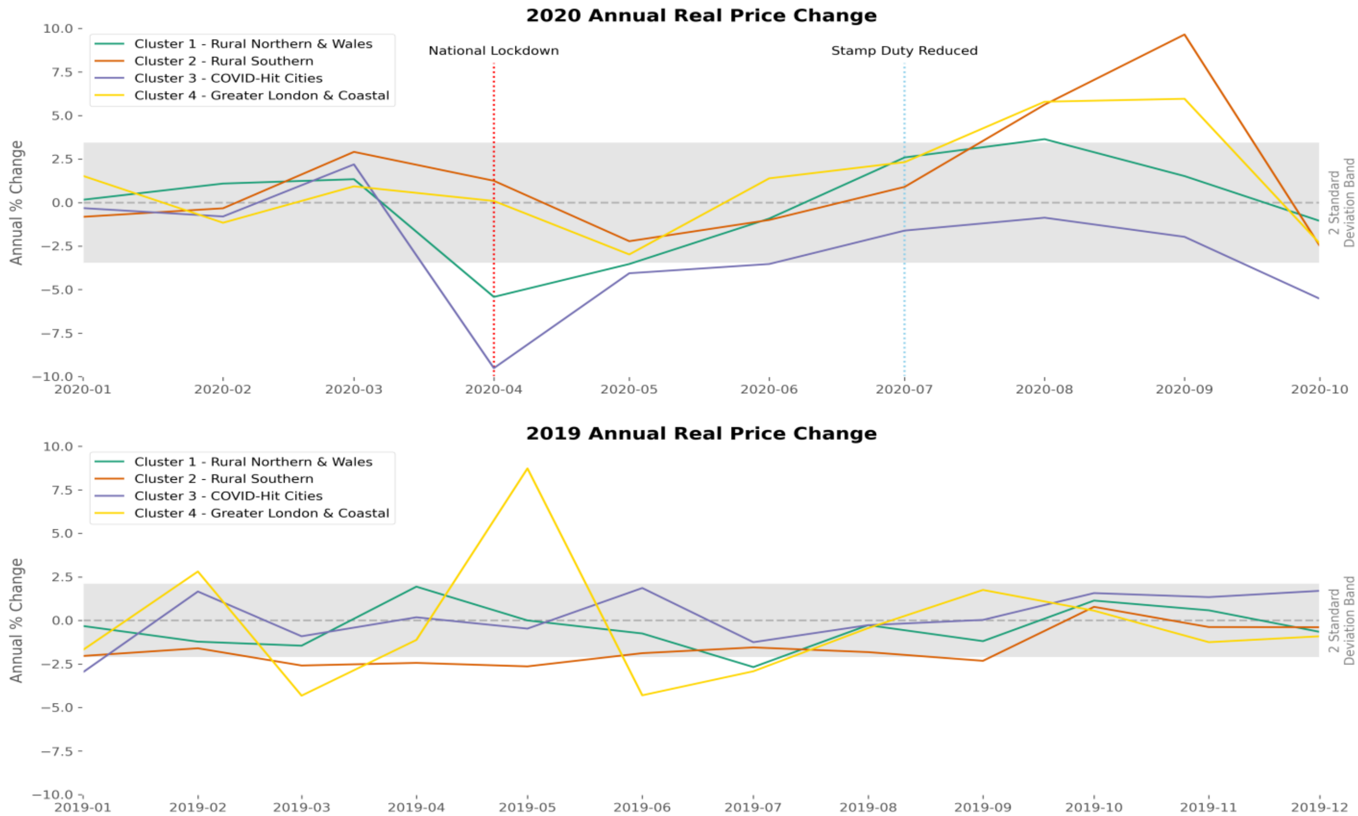


Fig. 10 – Time Series of Annual Real House Price Changes by Cluster

Analysis has revealed the exceptional nature of 2020, as house price changes overall have risen following a systemic shock whilst the number of houses being sold has fallen dramatically.

Clear temporal trends were identified during national and local lockdowns, after SDLT was reduced and as COVID-19 cases rose in different regions.

Spatial trends were identified as clustering produced four coherent cohorts which exhibited different price dynamics to each other, in contrast to 2019 price changes (Fig.10). Geographic heterogeneity was confirmed as an important factor by GWR's outperformance over MLR.

The regression model produced a good fit in comparison to the reference paper despite a lack of granular housing data, showing that publicly-available data can be used as a real-time indicator of house price reaction to the current pandemic.

These observations have meaningful implications for the government in evaluating the impact of the August stimulus, especially given this will end in March 2021. It is also relevant for future policy decisions as the government may need to provide further support for the hardest-hit counties included in cluster 3.

Furthermore, home buyers and sellers can be made aware of how COVID-19 has impacted house prices in their region and so can make more informed purchasing decisions.

5 Critical Reflection

Human reasoning was critical for the evolutionary process[30] of this study. Selecting relevant additional datasets to include and creating new features were processes that were largely dependent on my own prior knowledge combined with my interpretation of the visualizations constructed. Different perspectives could have provided additional inputs to this process, suggesting that, in order to make this analysis more robust, it should be carried about by a diverse team rather than one individual.

Refining the model inputs and choosing a value of k for clustering were also highly involved components of the research as optimization was essential to obtain a high value of R^2, and k had to be a small enough number so that any visual analysis could be easily understood, but large enough to ensure distinct groupings.

My research does not answer the research question related to the change in number of houses sold to the same extent as the annual house price changes. This was a conscious decision as I felt number of sales falling was more easily explainable, given the restrictions to movement during lockdown as well as the risk associated with house viewings. However, a deeper analysis of this factor could have revealed different spatio-temporal trends, either complimenting or counteracting my price change work.

Fig.9 shows that some of the cluster members overlap when visualized as PCA components. Fuzzy clustering[27], which allows datapoints to belong to more than one cluster, could be used as a comparative method. My approach assumes cluster membership remains constant over time, however fuzzy clustering could have enabled counties to shift between clusters throughout 2020 based on the changes to case numbers, or other metrics, in those areas.

In terms of software, utilizing tableau to construct time series and choropleth maps could have been more efficient than using python, potentially at the expense of control and customization.

The study was limited by access to housing data on a more granular level. Anyone wishing to build on this work may want to obtain access to this through the "Zoopla" developer API[28]. Whilst my work informs high-level takeaways that can be used in governmental policy-making, home-buyers rarely make decisions at a county-level and so analysis performed down to street or individual house level would be more valuable to this cohort. Using data at varying spatial scales could employ the recent extension to GWR – Multiscale GWR[28] – in order to fit a more complex model.

A further limitation of this work is that it is specific to the time period in focus – January 2020 to October 2020. There is no guarantee that the conclusions drawn can be extended throughout the COVID-19 pandemic as it progresses, and additional timely work is necessary to confirm whether the relationships found continue to hold. Because the nature of the shock is unique it also makes the inferences from this study less generalisable to shocks in the past or future. Furthermore, an extensive list of possible related variables was not sought and tested due to time constraints, so improved model performance and more accurate clustering could be possible.

Table of word counts

Problem statement	255
State of the art	510
Properties of the data	507
Analysis: Approach	511
Analysis: Process	1566
Analysis: Results	204
Critical reflection	508

References

The list below provides examples of formatting references.

[1] Goetzmann, W.N., Peng, L. & Yen, J. The Subprime Crisis and House Price Appreciation. J Real Estate Finan Econ 44, 36–66 (2012). https://doi.org/10.1007/s11146-011-9321-4

[2] https://www.gov.uk/guidance/stamp-duty-land-tax-temporary-reduced-rates

[3] Edwards, Michael. "The Housing Crisis and London." *City* 20, no. 2 (March 3, 2016): 222–37. https://doi.org/10.1080/13604813.2016.1145947.

[4] Holly, Sean, M. Hashem Pesaran, and Takashi Yamagata. "The Spatial and Temporal Diffusion of House Prices in the UK." *Journal of Urban Economics* 69, no. 1 (January 1, 2011): 2–23. https://doi.org/10.1016/j.jue.2010.08.002.

[5] Kang, Yuhao, Fan Zhang, Wenzhe Peng, Song Gao, Jinmeng Rao, Fabio Duarte, and Carlo Ratti. "Understanding House Price Appreciation Using Multi-Source Big Geo-Data and Machine Learning." *Land Use Policy*, July 21, 2020, 104919. https://doi.org/10.1016/j.landusepol.2020.104919.

[6] Brunsdon, Chris (09/03/2010). "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity". Geographical analysis (0016-7363), 28 (4), p. 281.

[7] Cao, Kai, Mi Diao, and Bo Wu. "A Big Data–Based Geographically Weighted Regression Model for Public Housing Prices: A Case Study in Singapore." *Annals of the American Association of Geographers* 109, no. 1 (January 2, 2019): 173–86. https://doi.org/10.1080/24694452.2018.1470925.

[8] Wu, Chao, Fu Ren, Wei Hu, and Qingyun Du. "Multiscale Geographically and Temporally Weighted Regression: Exploring the Spatiotemporal Determinants of Housing Prices." *International Journal of Geographical Information Science* 33, no. 3 (March 4, 2019): 489–511. https://doi.org/10.1080/13658816.2018.1545158.

[9] Waters, Nigel. "Tobler's First Law of Geography," December 12, 2017. https://doi.org/10.1002/9781118786352.wbieg1011.

[10] GOV.UK. "Price Paid Data." Accessed January 2, 2021. https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads. Contains HM Land Registry data © Crown copyright and database right 2020. This data is licensed under the Open Government Licence v3.0.

[11] "Official UK Coronavirus Dashboard." Accessed January 2, 2021. https://coronavirus.data.gov.uk/details/download.

[12] "Access to Gardens and Public Green Space in Great Britain - Office for National Statistics." Accessed January 2, 2021. https://www.ons.gov.uk/economy/environmentalaccounts/datasets/accesstogardensandpublicgreenspaceingreatbritain.

[13] AWS Ordnance Survey Blog. "Analysing Access to Greenspaces during a Pandemic," May 21, 2020. https://www.ordnancesurvey.co.uk/blog/2020/05/analysing-greenspace-access-during-pandemic/.

[14] "Population Estimates by Output Areas, Electoral, Health and Other Geographies QMI - Office for National Statistics." Accessed January 2, 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/smallareapopulationestimatesqmi.

[15] "RPI All Items Index: Jan 1987=100 - Office for National Statistics." Accessed January 2, 2021. https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/chaw/mm23.

[16] the Guardian. "UK's Most Expensive Home Valued at £160m," October 9, 2018. http://www.theguardian.com/uk-news/2018/oct/09/record-160m-paid-for-uks-most-expensive-home-ever-sold.

[17] Landler, Mark, and Benjamin Mueller. "In U.K.'s Test and Trace: Now You See 'em, Now You Don't." *The New York Times*, October 5, 2020, sec. World. https://www.nytimes.com/2020/10/05/world/europe/uk-testing-johnson-hancock.html.

[18] Andrew Reilly. "Buyer Demand Is Changing – Flats Less Attractive?" *The Good Estate Agent* (blog), July 16, 2020. https://thegoodestateagent.co.uk/buyer-demand-is-changing-flats-less-attractive/.

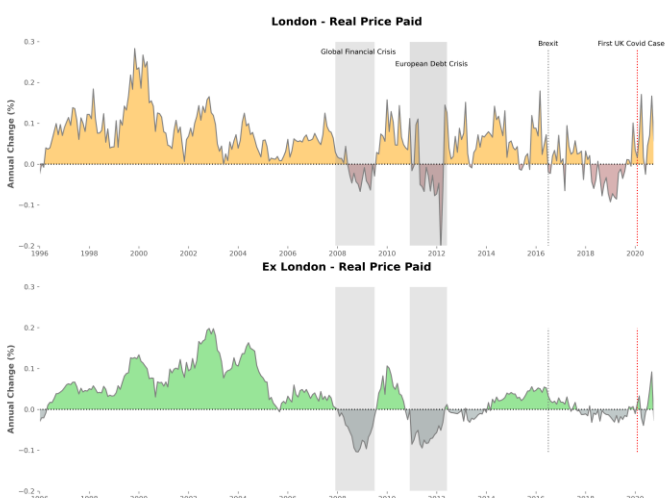
[19] the Guardian. "Covid Led to Huge London Property Exodus, Says Hamptons," December 26, 2020. http://www.theguardian.com/business/2020/dec/26/covid-led-to-huge-london-property-exodus-says-hamptons.

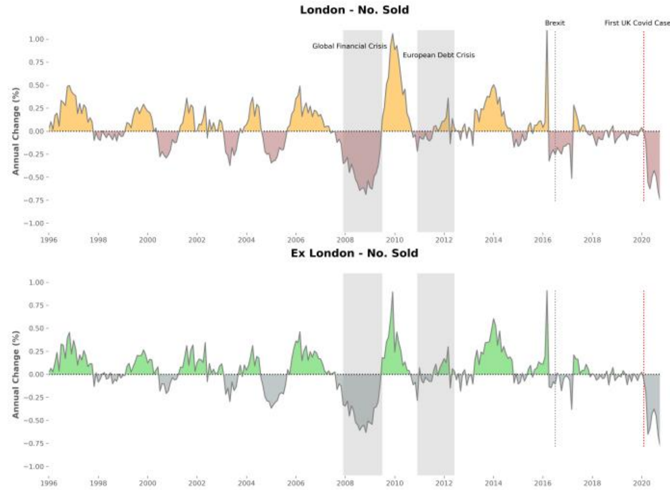
[20] Partington, Richard. "House Prices Will Drop in 2021 as Covid Impact Hits, Says Halifax." *The Guardian*, December 21, 2020, sec. Business. https://www.theguardian.com/business/2020/dec/21/house-prices-drop-2021-as-covid-impact-hits.

[21] GOV.UK. “UK House Price Index Summary: August 2020.” Accessed January 3, 2021. https://www.gov.uk/government/publications/uk-house-price-index-summary-august-2020/uk-house-price-index-summary-august-2020.
[22] “UK House Prices See Highest Growth in Six Years in 2020.” *BBC News*, December 30, 2020, sec. Business. https://www.bbc.com/news/business-55483432.
[23] Monmonier, Mark. “Strategies For The Visualization Of Geographic Time-Series Data.” *Cartographica: The International Journal for Geographic Information and Geovisualization* 27 (October 1, 1990): 30–45. https://doi.org/10.3138/U558-H737-6577-8U31.
[24] Charlton, Martin, Fotheringha, A, Stewart. “Geographically Weighted Regression White Paper”, National Centre for Geocomputation, National University of Ireland Maynooth, March 3, 2009.
[25] “Timeline of the COVID-19 Pandemic in the United Kingdom (January–June 2020).” In *Wikipedia*, January 2, 2021. https://en.wikipedia.org/w/index.php?title=Timeline_of_the_COVID-19_pandemic_in_the_United_Kingdom_(January%E2%80%93June_2020)&oldid=997823161.
[26] “The Challenges Facing Britain’s Housing Market after Brexit,” August 29, 2016. https://www.ft.com/content/aae1fa3e-6b7d-11e6-ae5b-a7cc5dd5a28c.
[27] Oliveira, Jose Valente de, and Witold Pedrycz. Advances in Fuzzy Clustering and Its Applications. John Wiley & Sons, 2007.
[28] “Zoopla Property API - Welcome to the Zoopla Developer Network.” Accessed January 5, 2021. https://developer.zoopla.co.uk/.
[29] Fotheringham, Alexander, Wenbai Yang, and Wei Kang. “Multiscale Geographically Weighted Regression (MGWR).” *Annals of the American Association of Geographers* 107 (August 28, 2017): 1247–65. https://doi.org/10.1080/24694452.2017.1352480.
[30] Andrienko, N., T. Lammarsch, G. Andrienko, G. Fuchs, D. Keim, S. Miksch, and A. Rind. “Viewing Visual Analytics as Model Building.” *Computer Graphics Forum* 37, no. 6 (2018): 275–99. https://doi.org/10.1111/cgf.13324.

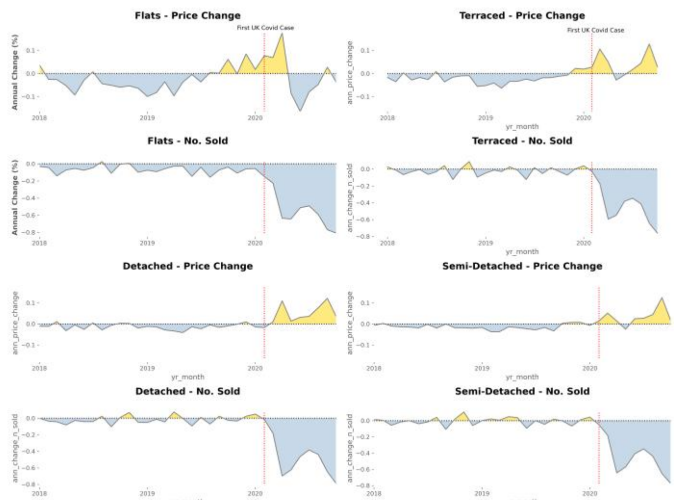
Appendices

Appendix A

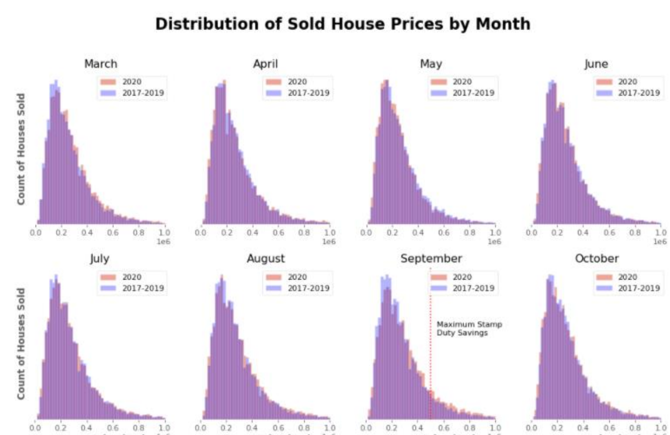


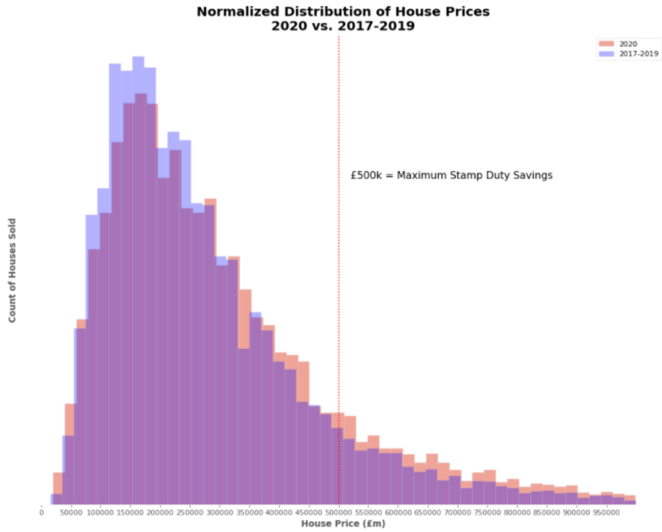


Appendix B

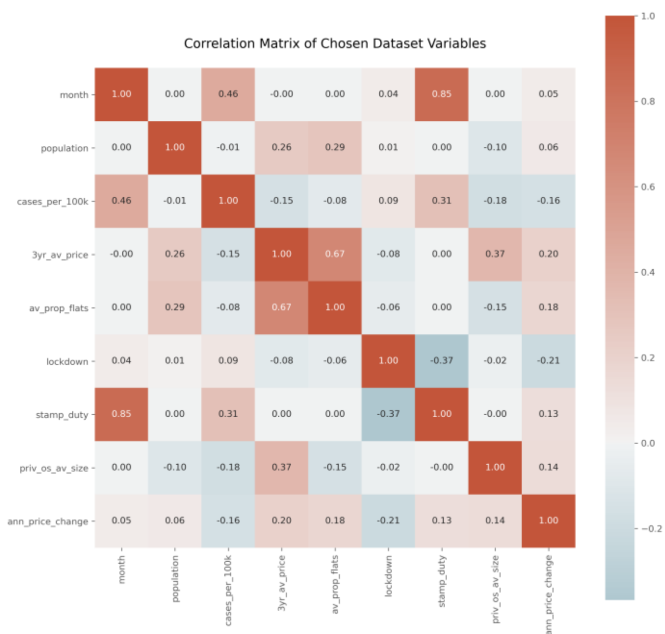


Appendix C





Appendix D



Appendix E

Summary of Cluster Variable Metrics:

Variables	Clusters			
	One	***Two***	***Three***	***Four***
3-Year Average Price (£)	225,522	360,744	157,091	317,603
Population	306,485	950,872	508,014	713,665
COVID-19 Cases Per 100k Pop.	125	77.5	260	118
Average Flat Proportion	8%	18%	8%	32%
Private Outdoor Space Size (Sq. m)	414	454	205	221

Clusters			
One – Rural Northern & Wales	***Two – Rural Southern***	***Three – COVID-Hit Cities***	***Four – Greater London & Coastal***
East Riding Of Yorkshire Gwynedd Staffordshire South Gloucestershire Cumbria Northumberland Denbighshire Cheshire West And Chester Ceredigion Pembrokeshire Flintshire Newport North Lincolnshire Isle Of Anglesey Swindon Northamptonshire Wrekin Bedford Shropshire Conwy Lincolnshire Powys Herefordshire Derbyshire Monmouthshire The Vale Of Glamorgan North Yorkshire Nottinghamshire Leicestershire Rutland Carmarthenshire Peterborough Cheshire East Torfaen Wrexham	North Somerset Wiltshire Isle Of Wight Buckinghamshi re West Sussex Norfolk Warwickshire Surrey West Berkshire Cornwall Hertfordshire Oxfordshire East Sussex Gloucestershire Cambridgeshire Worcestershire Wokingham Devon Kent Suffolk Bath And North East Somerset Central Bedfordshire Dorset Hampshire West Midlands Windsor And Maidenhead Somerset Essex	Merthyr Tydfil South Yorkshire Lancashire Greater Manchester Bridgend Middlesbrough Warrington Caerphilly Merseyside Blackburn With Darwen Halton Blaenau Gwent Tyne And Wear Leicester Nottingham Redcar And Cleveland Swansea West Yorkshire Hartlepool Darlington Kingston Upon Hull Rhondda Cynon Taff Blackpool Stockton-On-Tees Stoke-On-Trent County Durham Derby Neath Port Talbot North East Lincolnshire	Greater London Cardiff Portsmouth Luton Milton Keynes Southend-On-Sea Slough Medway Southampton York Thurrock Bournemouth, Christchurch And Poole Reading Bristol Plymouth Brighton And Hove Bracknell Forest Torbay