Assessment 3, Group 12

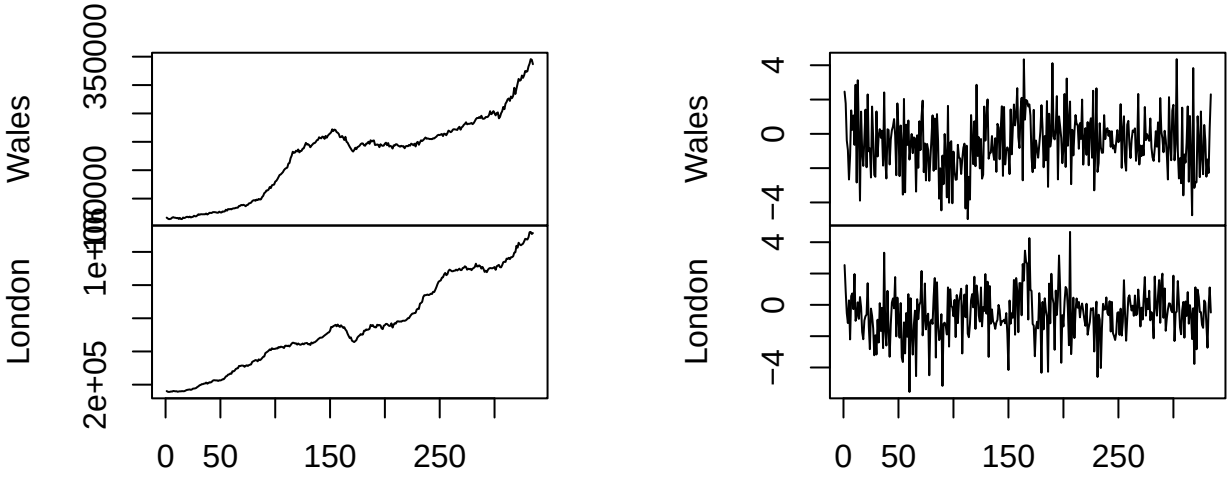
Joana Levtcheva (01252821), Sharon Salabi (2354789), Aidan van Niekerk (02363561)

UK Property Prices: An Exploratory Analysis

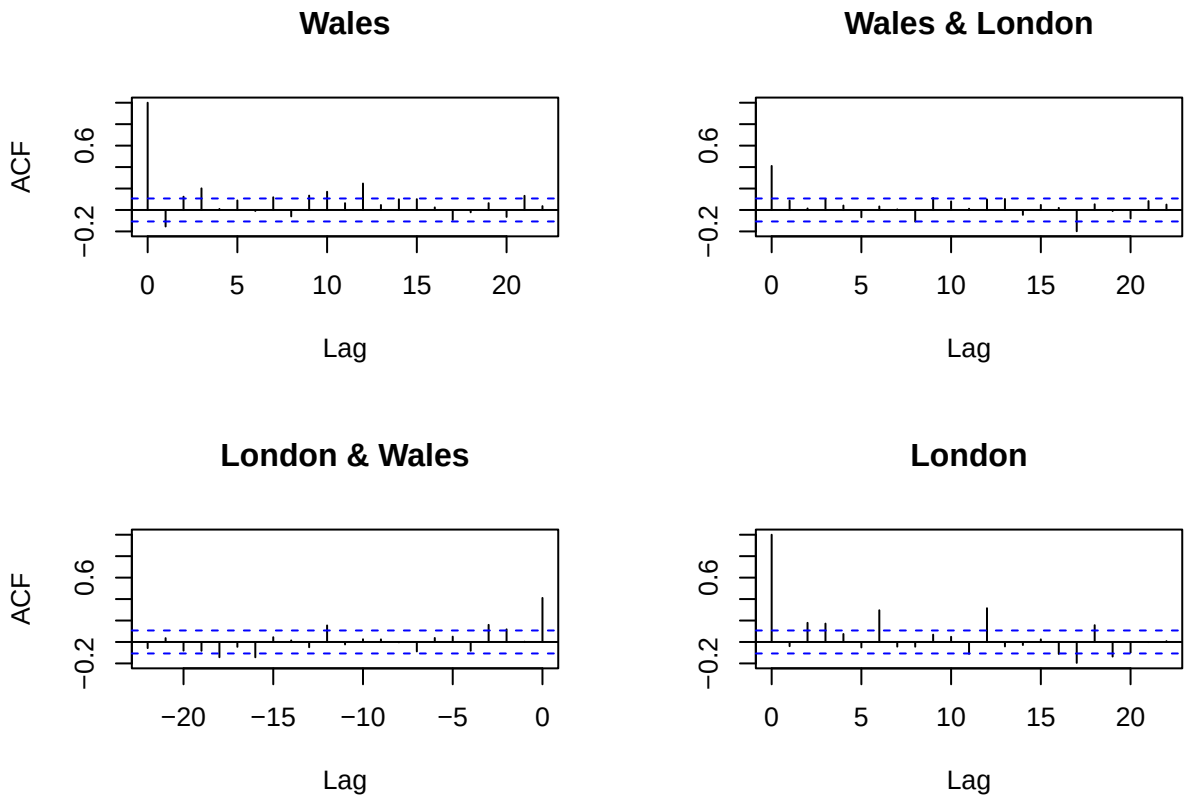
This report conducts a preliminary exploratory data analysis on the given UK property prices. The primary goals of this analysis was to examine the differences between two very different property markets - London and Wales. The two were chosen to compare and contrast a vibrant, densely-populated area with a more traditional and sparsely populated region. Specifically, a 15 year price percentage change was considered (coinciding with the financial crisis of 2007/2008). The effect of an highly economically active area such as London on nearby property valuations is discussed. Furthermore, property prices are examined during years of financial stress (such as 2008, Brexit, and COVID).

Multivariate Time Series EDA

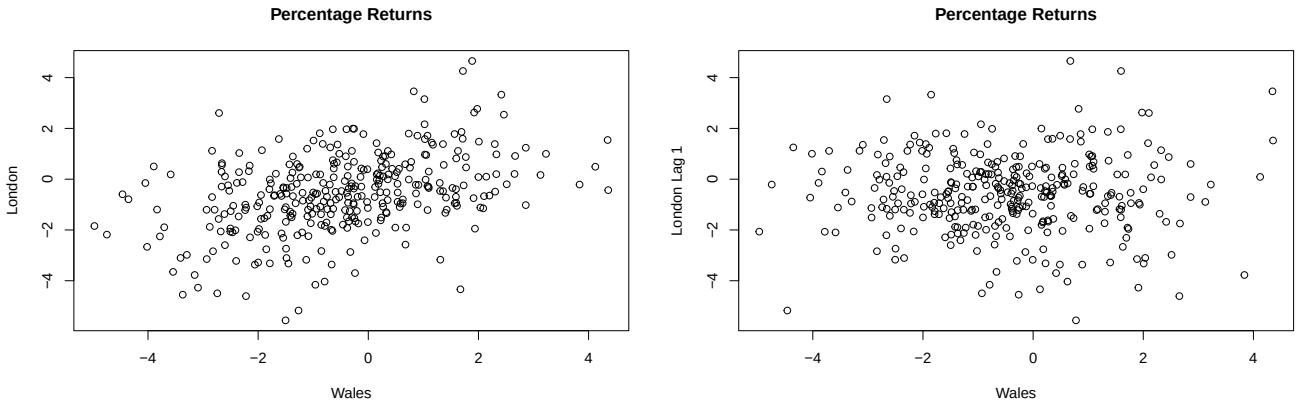
We are working with the monthly detached average prices for Wales and London, where there is no missing data and the observations for both regions start and end on the same date. In order to model the prices (left figure) as a stationary bivariate time series, we first reexpress the data as percentage relative price changes (right figure) by $X_{t1} = 100\frac{W_t - W_{t-1}}{W_{t-1}}$ and $X_{t2} = 100\frac{L_t - L_{t-1}}{L_{t-1}}$. We are observing the bivariate time series $X_t = (X_{t1}, X_{t2})'$, where X_{t1} corresponds to Wales data and X_{t2} to London.



From the sample autocorrelations and cross-correlations we see that the autocorrelations ρ_{11} (Wales) and ρ_{22} (London) are generally small, there is also a correlation between $X_{t,1}$ and $X_{t-1,2}$. This suggests the London relative price change on month $t-1$ may be of assistance in predicting the Wales relative price change on month t.



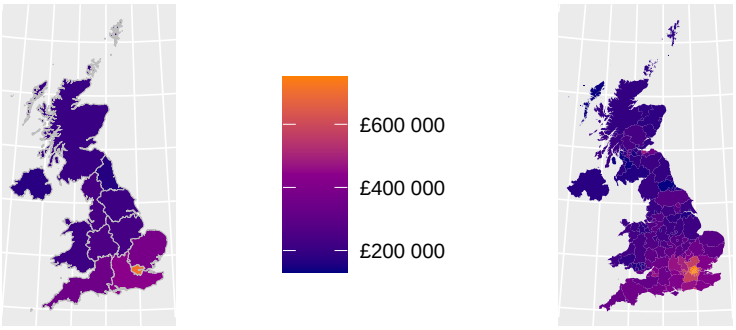
This can be supported by the scatterplot on the right of the points $(x_{t-1,1}, x_{t,2})$, $t = 2, ..., 334$:



When testing for Zero Cross-Correlations with the multivariate Portmanteau test, the p-values are equal to 0 so we reject the null hypothesis. This confirms the existence of linear dynamic dependence in the bivariate return series at the 5% significance level.

Spatial EDA

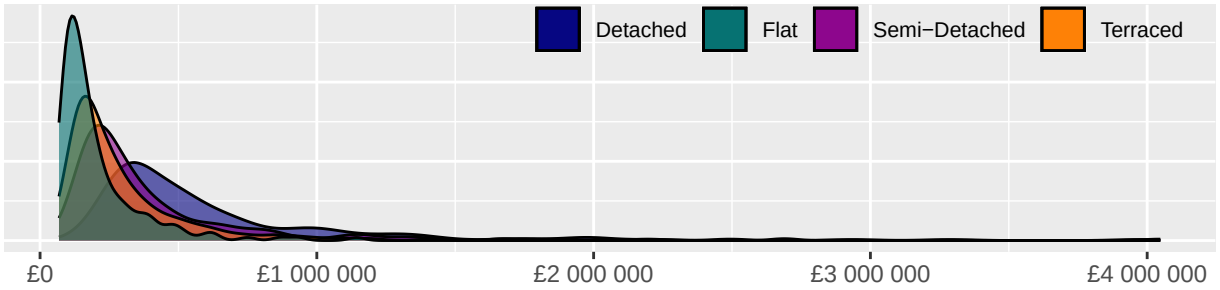
In order to gauge a rough idea of the data, two simple showing the semi-detached average of all property types across a) ITL 2 and b) ITL 3 levels are presented. Semi-detached prices tend to be in the middle of the range of property types and is thus a useful metric in ascertaining an initial impression.



Both levels of the data, regional and county, show a clear clustering: higher property prices appear to be associated with closer proximity to central London. To a lesser extent, there appears to be a "hotspot" on Edinburgh evident on the right plot. Northern Ireland tends to have lower property prices evident on both a regional and county level with no "hotspot" evident for the areas surrounding Belfast. For the purpose of EDA, ITL 3 level will be considered.

Distributional Properties

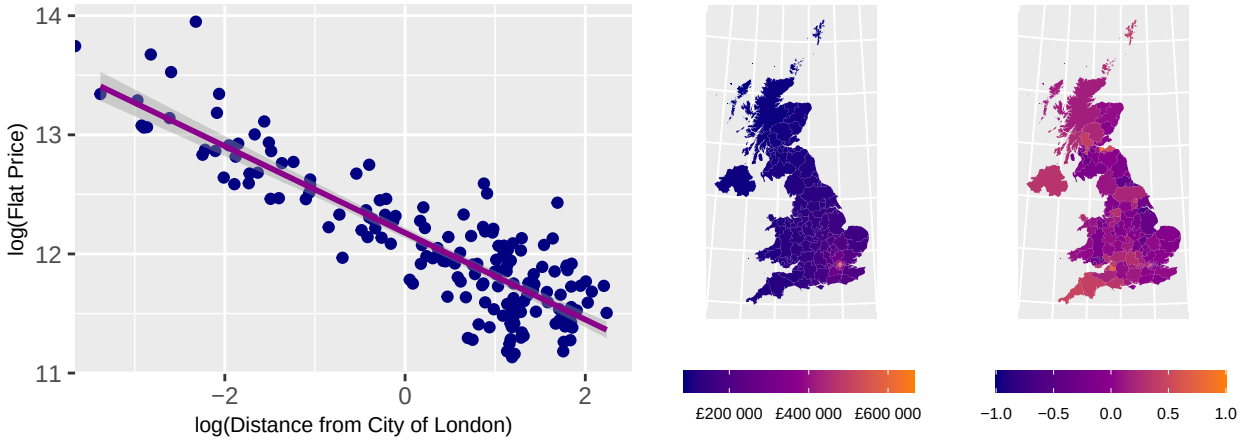
The kernel densities below by property type clearly shows that the distribution of the various types are different. The table of summary statistics supports this claim as the means, medians, and standard deviations are substantially different. On average, detached prices are higher than semi-detached, terraced, and flat prices (in that order). Their standard deviations are also of the same order - indicating more expensive properties types have wider dispersion around their means. All types exhibit a fairly large positive skew - but interestingly semi-detached has the greatest maximum value of all types. Missing values are few and not of concern for later analysis, however it is noted that City of London only has a flat property value. The implication is that regions that are typically more rural (further afield from cities) are associated with having lower flat prices.



	Min	Max	Lower Q	Upper Q	IQR	Median	Mean	SD	N	NAs
Detached	208 593	3 968 212	327 253	693 166	365 913	458 102	623 757	531 181	176	2
Semi−Detached	134 115	4 047 465	192 931	461 750	268 819	278 212	411 313	431 976	176	1
Terraced	99 981	2 689 282	157 752	360 487	202 735	222 723	316 821	301 317	176	1
Flat	68 415	1 143 867	107 812	228 648	120 836	150 081	206 040	161 108	176	0

Trend component: all roads lead to London

To account for some variation in the spatial data, we examine the relationship between the logarithm of flat prices and the logarithm of distance from the City of London. The left plot clearly indicates a negative linear association between the aforementioned variables. When a linear model is fit with log(Flat Price) as a response and log(Distance) as a predictor, a plot of the fitted values is produced (middle). As expected, it displays a pattern of exponential decrease as one gets further from City of London. It provides a decent approximation when compared to the plot of the actual property values (above). Examination of the residual plot (right) shows a plot consistent with random residuals. It is noted, however, that the model has positive residuals near cities such as Bristol and London where it under estimates property prices. Generally speaking, the described linear model does a decent job of accounting for variation of flat prices when accounting for distance from City of London.

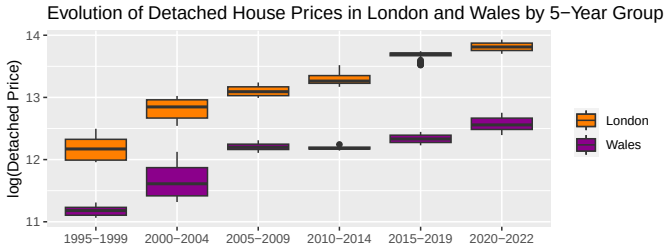


Visualisation

In this section, we explore additional visualizations that provide deeper insights into our dataset and showcase it from a different perspective.

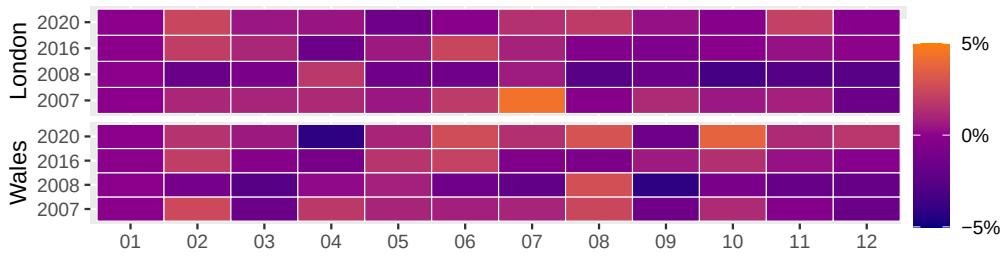
Detached Property Prices by 5 Year Group

A linear rise in the average price of detached properties in London can be observed from the boxplot data grouped in 5-year intervals. Additionally, it can be noted that the price variability has decreased over time, indicating that the cost of detached properties in London has increased while the range of prices has become narrower. In contrast, Wales experienced a significant increase in detached property prices between 1995 and 2009, followed by a slight decline, and a slower subsequent increase. The range of prices for detached properties in Wales is generally narrow, with an exception of the period between 2000 and 2004, during which prices were more widely dispersed.



Monthly Percentage Change: 2007/2008 (Financial Crisis), 2016 (Brexit), 2020 (COVID)

The plot illustrates the monthly percentage change in detached property prices in London and Wales during specific crisis years. Notable differences can be observed in how the two locations responded to these crises. For instance, during the financial crisis in July 2007, London saw a 5% increase in detached property prices, whereas Wales did not experience a similar increase. In September 2008, the average prices in Wales dropped by 5%, which was a more significant decline than that seen in London. During the COVID-19 crisis, London's detached property prices remained stagnant, while Wales experienced more price fluctuations. In April 2020, a month after the border closure, prices dropped in Wales and did not recover until October. Interestingly, Brexit did not have a significant impact on detached property prices in either London or Wales.



Highest and Lowest Changes over 15 Years

This visualization displays the maximum and minimum percentage changes for each type of property across different regions and countries of the United Kingdom in the past 15 years. Notably, detached properties, despite being the most expensive type of property, experienced the highest increase in most regions of the UK, with the exception of London, where flats had the highest increase, and the South West, where terraced properties saw the highest change. The second map displays the minimum percentage change for each type of property in different regions and countries of the United Kingdom over the past 15 years. It is evident that semi-detached properties had the least change overall, except in the North East region where terraced properties experienced the lowest change.