Analysis of the London Housing Market

Final Project - Data Stream - Group 6

Ekaterina Irimca | Katherine Cupit | Maria Celeste Orona| Monica Escobar Cuesta | Nasra Elmi

---

1.   INTRODUCTION

The London housing market plays a pivotal role in the city's economy and attracts significant attention from investors and real estate professionals. This project aims to analyse the evolution of property prices in London over time and identify key factors that influence these trends.

Questions we will execute:
1. How does the housing market in London react to major events such as the 2008 Financial Crisis, Covid-19, Brexit and key economic indicators such as inflation and affordability?
2. Do property price trends affect every region within London in the same way and how do these trends compare to the rest of Europe? How did the demand evolve during these years? Are certain regions or types of properties becoming more desirable than others?
3. How does the number of new dwellings in London impact house prices?
4. To what extent do Housing Supply Indicators, Housing Demand Indicators, and Average Price interrelate within the London Market?

Objectives:
- Identify insights into the housing market's resilience and sensitivity to macroeconomic factors and understand which external events are the most impactful on London house prices.
- Identify any disparities in house prices across London and potential hotspots.
- Establish the relationship between housing supply, demand and prices in London. Provide insights into housing affordability, areas of potential housing shortages, and the effect construction of new dwellings has on the housing market.

1.1. Roadmap of the project

Week 1 (28.07-03.07): Data collection, cleaning and preprocessing. Identify any inconsistencies or missing values.

Week 2 (03.08-10.08): Analyse the clean data, calculate key metrics and create visualisations.

Week 3 (10.08-17.08): Identifying influencing factors and relationships between variables. Research and analyse factors influencing the UK housing market, e.g., inflation, base rates, affordability. Look at the relationships between housing supply, housing demand and prices.

Week 4 (17.08-24.08): Draw conclusions and summarise findings. Compile all analysis performed and prepare the information for presentation.

2. BACKGROUND

The London housing market is a key driver of economic activity and wealth creation within the city. The value of properties, changes in property prices, and the affordability of housing directly impact various facets of life, including financial stability, investment opportunities, and overall quality of living.

This project delves into the intricacies of the London housing market and explores the evolution of property prices over time, aiming to uncover key insights and trends that shape this critical sector. By examining the reactions of the housing market to significant events such as the 2008 Financial Crisis and the Covid-19 pandemic, the project seeks to shed light on the market's resilience and vulnerability to external shocks.

We expect our findings to hold significant implications for a diverse audience. Real estate professionals can leverage the insights to make informed decisions regarding property investments, while policymakers can

gain a deeper understanding of the housing market's response to economic shifts and policy changes. Additionally, potential homebuyers and sellers can use the project's insights to navigate the market and make well-informed decisions regarding property transactions.

3. SPECIFICATIONS AND DESIGN

3.1 Design Flowchart
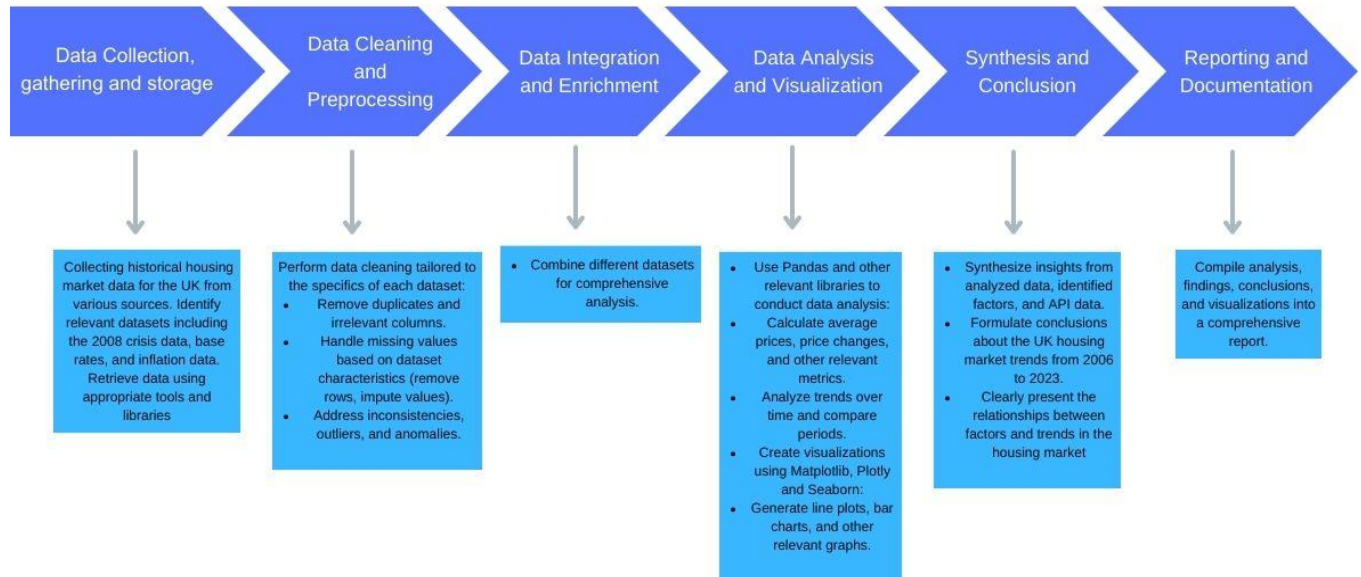
3.2 Technical Requirements

- Collect relevant data to the chosen topic.
- Pre-process and cleanse all of the datasets.
- Analyse data and visualise findings utilising Pandas, Matplotlib, Plotly and Seaborn libraries.

3.3 Non-Technical Requirements

- Familiarity with the subject of interest.
- Prior knowledge of how to use APIs to access data.
- Ability to interpret results gained through our analysis.
- Communication and collaboration skills.

4. IMPLEMENTATION AND EXECUTION

4.1 Development approach and team member roles

- Agile approach with biweekly meetings (more details in 4.3. below)

4.2 Tools and libraries

- Matplotlib and Seaborn for creating diverse visualisations, including line plots, bar charts, and heatmaps
- Pandas for data manipulation, preprocessing, and exploratory analysis.
- Scipy.stats for working out correlation.
- Datetime for handling and manipulation of date-based data
- Json for using GeoJSON files in choropleth maps.
- Plotly for plotting the choropleth map.
- Eurostat Python package for retrieving and filtering data from Eurostat.
- Bar_chart_race for creating EU HPI comparison bar chart race.

4.3 Implementation process (achievements, challenges, decision to change something)

<u>**Changes**</u>

Initially, we decided to use GitHub. However, some of our data sheets were too heavy to push and pull comfortably to and from GitHub. This is why <u>we decided to use a shared drive instead</u>. This decision was taken as a team, after making sure that everyone agreed to the change.

<u>**Challenges**</u>

As a *<u>challenge</u>*, every team member has different schedules and this has made it sometimes difficult to have meetings where all the members were present. However, by keeping the minutes of every meeting we have been able to maintain the sense of group and allow for everyone to take part.
We created a shared drive folder, so we could all work asynchronously.

One of the initial challenges that we faced was finding a relevant dataset accessible through an API. After exploring different options, we decided to collect House Price Index information from Eurostat, the statistical office of the European Union.

<u>**Achievements**</u>

There has been clear interest in the project by all members from the beginning. This has helped enormously to work as a team and to produce quality work where feedback was given by all members of the group and no one doubted to step in to improve something or to support, this has positively impacted our timeline and the final analysis. We have all felt supported by the rest of the team members and all our roles were clear, some of the team members focused more on the cleaning part and some others were more keen on working on the visualisation section, even though we have all worked together.

We have ensured consistent and accurate data by using various sources.

We followed an Agile methodology, which allowed for:

- Iterative Approach: We regularly reviewed and refined analysis techniques, visualisation strategies, and interpretations based on emerging insights.
- Refactoring: We optimised code and visualisations iteratively to enhance clarity, efficiency, and user experience.
- Code Reviews: Regularly did peer code reviews to facilitate quality assurance and knowledge sharing among team members.

4.4. Data Collection

Information needed (from 2006 to 2023)	Data Source/ Methodology
Historical Property Price Data. This dataset includes historical property sale prices, and property attributes (such as type, size, location) for various areas within London. This data will allow us to track the changes in property prices over time and understand how different factors influence these changes.	- HM Land Registry Open Data - London Datastore (UK House Price Index.xlsx)
Economic and Demographic Indicators (inflation and interest rates).	- Bank Of England (base_rates.xls) - ons.gov.uk (CPIH_annual_rates.csv) - ons.gov.uk (Affordability.xlsx) - Ons.gov.uk (CPIannualrate.xls)
Geospatial Data: To analyse how property prices vary across different geographical areas in London, the rest of the UK and Europe.	- Eurostat API - London Datastore (london_boroughs.geojson)

General Information on significant events that might have influenced the housing market, such as financial crises or COVID-19.	- General information available on COVID-19, 2008 financial crisis and Brexit dates obtained through official online sources.
Data on Housing Market Dynamics: to determine relationships between Housing Supply Indicators (New and Total), Demand Indicators (New and Total) and Average Price (for New Builds and All Properties).	- London Datastore: Total Number of Dwellings and Net Additional Dwellings, Borough. - Live tables on Housing Supply: Indicators of new supply. (LiveTable253a.xlsx) - Land Registry all UKHPI data for Inner, Outer and City of London regions.
Data collection through Eurostat.	The information from Eurostat can be accessed programmatically through their API, which provides data in JSON-stat 2.0 format and supports REST protocol. When making a query through the API, the response is a deeply nested json object, which is complex to read and filter. Therefore, to obtain the data we decided to use the Eurostat Python Package, which is a tool that uses Eurostat SDMX 2.1 API web services to retrieve, filter and convert the information into a Pandas DataFrame.

5. RESULT REPORTING AND CONCLUSIONS

1. _How does the housing market in London react to major events such as the 2008 Financial Crisis, Covid-19, Brexit and key economic indicators such as inflation or affordability?_
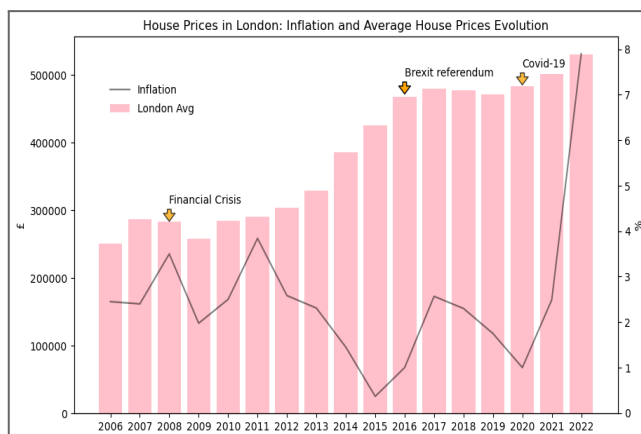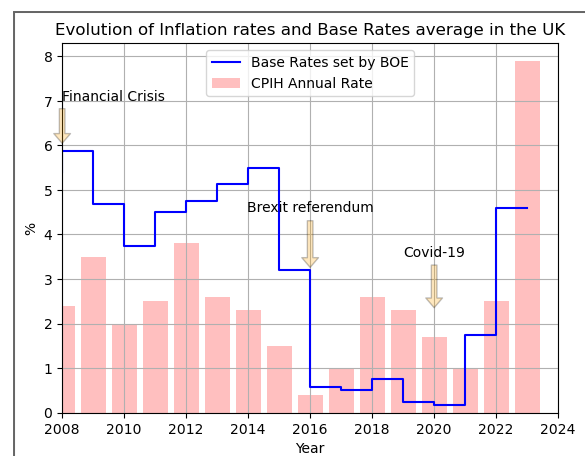

Figure 1- Inflation and house prices average in the UK.

Inflation is the rate at which the costs of commodities and services surge over time. For the analysis purposes of this project, the CPIH has been used as it covers the index of consumer goods and services and costs associated with owning and maintaining one's property.

The **correlation** between inflation and house prices changes in London is 0.074. This means that house prices are not affected by inflation and it is easy to see in figure 1 how house prices have increased gradually over time without being directly affected by inflation, showing a weak correlation between the inflation and house prices, as stated above by the correlation figure.

We can also see how the Brexit referendum did not have a significant impact on house prices, unlike the financial crisis in 2008. During this period, which started in 2008, thousands of citizens lost their jobs, and/or were unemployed. This had a clear negative impact on house prices which fell by 9.2 % in 2009. House prices started their slow recovery in 2010, progressively increasing house prices until 2023 (2018 and 2019 had a slight dip that was a reflection of the uncertainty around the Brexit vote, however, it's not statistically significant).

Figure 2- Evolution of base rates and inflation

COVID-19 had a strong impact on the housing market. As working from home became the norm, along with a Stamp Duty relief period and record low base interest rates set by the Bank Of England, housing demand increased steeply as soon as the housing market was back open after lockdown.
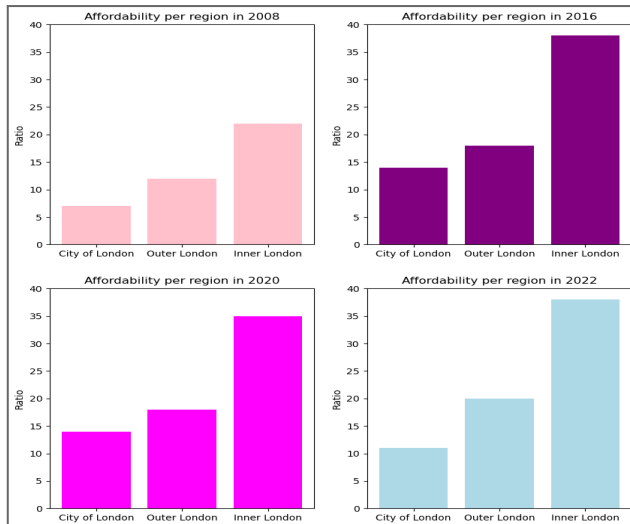
Figure 3- Affordability per region in London.

Affordability is measured as the percentage of average earnings needed to afford a mortgage on a typical home in each borough and is linked to the cost of living and therefore, inflation. This makes it a key factor to analyse as one of the key indicators for this report to gather conclusions on how the housing market is affected by different key economic indicators.

The affordability of housing in London varies significantly across the boroughs. After analysing the data, we can safely say that the most affordable boroughs are Barking and Dagenham, Bexley, and Bromley, whereas the least affordable boroughs are Westminster, Kensington and Chelsea, Hammersmith and Fulham (see Appendix).

In figure 3, we can see how affordability has evolved from 2008, adding more financial pressure onto household finances overtime.

The decline in affordability is due to a number of factors, including the rising cost of housing, the current stagnant wages and increase in the cost of living as a result of increased inflation.

We can see how in some boroughs, especially the ones in Inner London, the affordability ratio is higher now than it was in 2020 (COVID-19 period), meaning that these boroughs have become less affordable over time, except for the City of London.
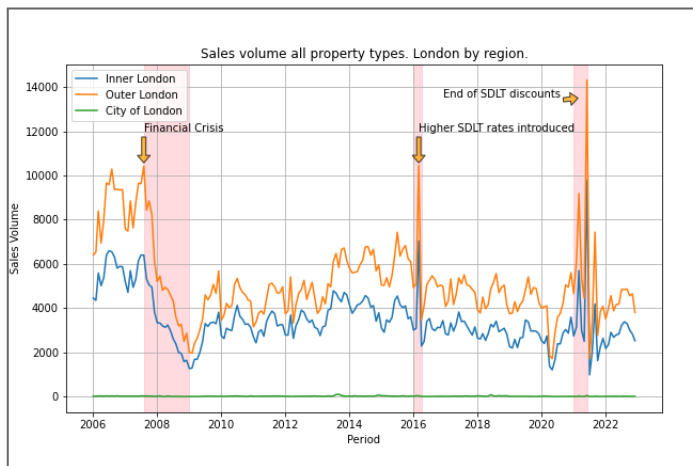

Figure 4- Sales volume (all regions in London).

As for the City of London we can conclude that it is more affordable than the rest of London. A few of the reasons are that the City of London is mostly office space, not a residential area and offers high income jobs.

Overall, the City of London is more affordable than Outer London because it has a lower demand for housing (based on sales) and a higher average income, the two factors taken into account when working out affordability.

When analysing all of the report figures we can come to the conclusion that the Brexit referendum and the COVID pandemic did not have a significant impact on house prices, unlike the financial crisis in 2008 and the current period.

Inflation and the affordability ratio were both lower in 2008 compared to 2022. We can see how the market is behaving differently in both scenarios as we can see how house sales were going up in 2022 for Inner and Outer London. We are starting to see a gradual decline in 2023 but it's too early to draw any conclusions yet.

2. *Do property price trends affect every region within London in the same way and how do these trends compare to the rest of Europe and the UK?*

Overall, the House Price Index in London follows the general trend of the HPI in the rest of the UK, however, the average price paid in London remained significantly higher than the rest of the UK. During the period analysed, we can see the decline of prices in all regions due to the 2008 financial crisis, while they slowly started to recover during the second half of 2009, which corresponds to the moment when the UK officially came out of recession

(last quarter of 2009). . However, it took the UK's GDP 5 years to recover to pre-crisis levels, which finally happened in September 2013. At this point, we also see the start of a period of rapid growth in the house price index.

Until 2020, we see no significant differences in the HPI between Inner and Outer London, however, since that year we can see that prices in Outer London started to increase at a faster rate.

From the start of the Covid-19 pandemic in March 2020 to September 2022, the average price paid in Inner London increased by 9.1%, while in Outer London, it increased by 16.4%. This is due to the relative affordability of Outer London compared to Inner London, which means that as prices rise, buyers are pushed further out of the centre.

Additionally, the adoption of remote work during the Covid-19 pandemic, means that people do not mind living further away.
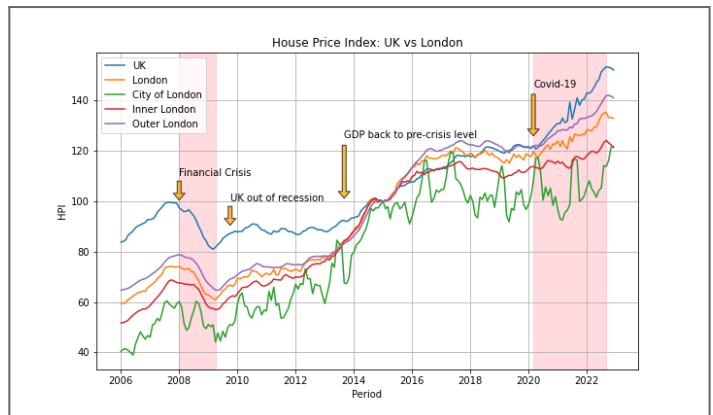

Figure 5- HPI UK vs London

Average prices December 2022:

UK: £289956.
London: £535407.
Inner London: £622388.
Outer London: £491926.
City of London: £953985

When comparing the HPI of London and the UK to the rest of Europe, we can see that on average, the evolution of house prices in the UK follows the trends of the rest of Europe.
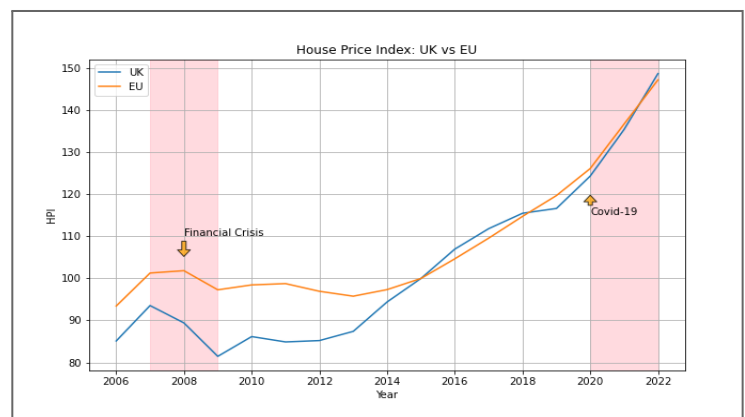

Figure 6- HPI UK and EU.

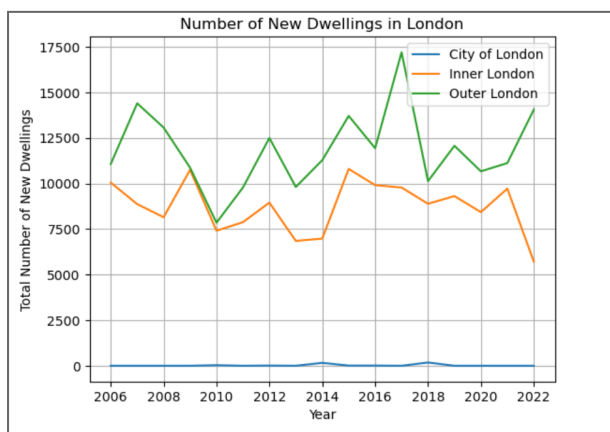3. _How does the number of new dwellings in London impact house prices?_

Figure 7- Number of new dwellings in London.

Figure 7 shows outer London has the highest number of total new dwellings each year from 2006 to 2022, which may be due to lower building costs and more available space than in the City of London and inner London. There has been a further increase in outer London since COVID-19 in 2021 with more people moving to this region from inner London, as they were not having to commute therefore the demand for new dwellings in inner London has decreased. The City of London is an outlier in this case, maybe due to empty land for new dwellings being so scarce here. However, this reinforces the theory that a lower supply of new dwellings increases the average house price, as the City of London is known for its expensive housing.

Figure 8 shows Tower Hamlets is the borough with the overall largest increase in total dwelling stock between 2006 and 2021. This suggests that Tower Hamlets is a
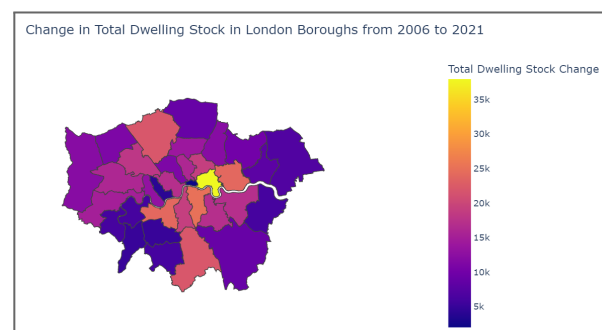

Figure 8- Change in total dwelling stock per borough.

hotspot in London for new builds due to a demand for affordable housing and a high population density. Economically well-off boroughs, such as Kensington and Chelsea and Kingston upon Thames, saw a very small increase in their total dwelling stock, which may be due to a lower demand for flats and affordable housing as residents in this borough generally have higher incomes and chose to live in larger, more expensive houses, therefore there is a smaller number of houses.

Figure 9 shows that Kensington and Chelsea had the highest average house prices at the end of 2021 (around £1.4 million), which aligns with the previous choropleth map (Figure 8) showing this borough to have a very small increase in their total dwelling stock. This reinforces the hypothesis that areas with a low number of new dwellings tend to have higher house prices as there is a lower supply and a higher demand. Although Tower Hamlets has a relatively low average house price in London at around £440,000, Barking and Dagenham is the lowest at around £331,000, which does not align with Figure 8 as Barking and Dagenham had a very low total stock.
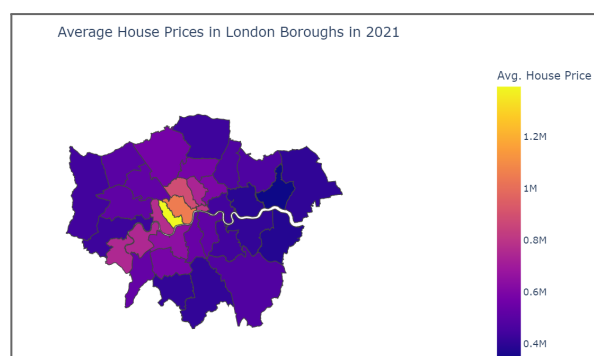

Figure 9- Average house prices per borough.

In conclusion, our analysis shows that there is a weak pattern between the supply of new dwellings and average house prices in London, as boroughs with a low total dwelling stock tend to have higher average house prices than those that have a high total dwelling stock. However, there are many other factors that play a part in an area's average house price, such as the local schools, the average size of the properties and local amenities.

4. *To what extent do Housing Supply Indicators, Housing Demand Indicators, and Average Price interrelate within the London Market?*

Relationship between:	Correlation Coefficient for each region:	Findings and conclusions:
Completed New Dwellings and Total Sales for New Builds *(For the visualisation illustrating the relationship between variables, please refer to Figure 1 in Correlations.pdf)*	Inner London: 0.5 Outer London: 0.3 City of London: 0.6	The correlation coefficient is highest for the City of London, followed by Inner London and then Outer London. This suggests that for the City of London and Inner London, completed new dwellings have a moderate positive relationship with total sales for new builds. Possible reason for this is strong demand for newly constructed properties. For Outer London the correlation between new dwellings completed and sales for new builds is weak, meaning that variables move in the same direction only in certain instances.
Average Price for New Builds and Completed New Dwellings *(For the visualisation illustrating the relationship between variables, please refer to Figure 2 in Correlations.pdf.)*	Inner London: -0.03 Outer London: 0.3 City of London: 0.2 (up to 2021)	The correlation coefficients indicate mixed relationships between completed new dwellings and average price for new builds. Outer London has the strongest positive correlation, however still weak. This indicates that in some cases as the average price for new builds tends to rise, completed new dwellings tend to increase as well. One possible explanation for this could be that construction providers see an opportunity for higher profit margins. For the Inner London Region, there's minimal consistent evidence suggesting a link between fluctuations in the two variables, due to the correlation of -0.03 being not significant.
Total Dwellings Stock and Total Sales for All Properties (up to 2021) *(For the visualisation illustrating the relationship between variables, please*	Inner London: -0.5 Outer London: -0.4 City of London: -0.2	All regions show negative correlations, Inner London has the highest negative correlation, suggesting that there is a moderately negative relationship between total dwellings stock and all property sales. Outer London has slightly lower negative correlation and City of London has the weakest one out of them all. For Inner and Outer London, this relationship might be influenced by an oversupply situation leading to reduced demand. However, for the City of London, this may not be as applicable due to weaker correlation.

refer to Figure 3 in Correlations.pdf):		
Average Price and Total Dwellings Stock for All Properties (up to 2021) *(For the visualisation illustrating the relationship between variables, please refer to Figure 4 in Correlations.pdf.)*	Inner London: 0.9383 Outer London: 0.9396 City of London: 0.8	All regions have strong positive correlations, indicating that as the average property prices increase, the amount of total dwellings stock available increases as well. Outer and Inner London have slightly higher and very similar correlations. For these regions, the relationship can be attributed to a slightly stronger demand driven by the availability of more affordable properties, thereby leading to an increase in property development and property values. As for the City of London, the relationship can be caused by the elevated demand resulting from the region's status as a financial hub that consistently attracts professionals and investors from around the world.
Average Price and Total Sales for All Properties (up to 2021) *(For the visualisation illustrating the relationship between variables, please refer to Figure 5 in Correlations.pdf.)*	Inner London: -0.4 Outer London: -0.2 City of London: -0.09	Inner and Outer London regions have moderate and weak negative correlations, respectively. This indicates that in several instances (only for some cases in Outer London), when the average house price increases, the total sales for all properties decrease. The relationship may be caused by potential buyers becoming more cautious due to an increase in price or affordability problems, leading to a decrease in overall property sales. For the City of London the average price variable does not have any connection with total sales, as the negative correlation of -0.09 is not significant.

Conclusion: Correlations provide the insights into the relationships between variables, however, do not imply causation. Analysis above highlights potential explanations for the interplay between variables, other factors can also influence the demand and supply dynamics in the housing market.