

Politecnico di Torino

Master of Science in

Territorial, Urban, Environmental, and Landscape Planning

2026

The Influence of Urban Green Amenities on the Housing Market: An Application to the City of London

Tutor

Luigi Buzzacchi

Candidate

Giovanna Padovan Garces

Words cannot express my gratitude to my tutor for his invaluable patience and feedback.
I am also grateful to my classmates and friends who supported and motivated me during this process.
Not to mention my parents. Their belief in me is what made it all possible.
Lastly, I could not have undertaken this journey without my best friend and lover, William, for his editing help,
late-night coffee supply, and moral support.

EXECUTIVE SUMMARY

This study investigates the extent to which urban green infrastructure is capitalised into residential property values in London, with particular emphasis on the spatially heterogeneous effects of proximity to major green spaces. Drawing upon the hedonic pricing framework, the analysis estimates the implicit value of green amenities through a series of regression models in which inflation-adjusted residential property prices per square metre are explained by a range of structural, building-quality, and locational characteristics, including property type, floor area, construction period, energy performance, glazing characteristics, accessibility to transport and educational facilities, and distance to substantial green spaces. The empirical results demonstrate that the incorporation of locational and neighbourhood characteristics substantially improves model explanatory power, while the inclusion of green-space proximity produces a comparatively modest but statistically significant contribution to the explanation of housing-price variation. District-level estimates further reveal considerable spatial heterogeneity: proximity to green infrastructure is positively capitalised into property values in selected districts, particularly where accessible green space constitutes a relatively scarce residential amenity, whereas its effect is weaker or statistically insignificant in highly central and spatially heterogeneous areas where other locational attributes exert greater influence. These findings indicate that the economic valuation of urban green infrastructure cannot be adequately understood through distance-based measures alone, as the magnitude and direction of its association with property values are contingent upon the broader spatial, socioeconomic, and market context in which green amenities are embedded. The study consequently underscores the relevance of context-sensitive approaches to environmental valuation and urban planning, while recognising methodological limitations arising from omitted housing characteristics, the operationalisation of green accessibility through proximity, and the geographical specificity of the analysis.

Keywords: *Urban Green Infrastructure, Hedonic Pricing Model, Residential Property Values, Environmental Amenities, Spatial Heterogeneity, Housing Market, London.*

LIST OF ABBREVIATIONS

ES	Ecosystem Services
GI	Green Infrastructure
UCL	University College London
GLA	Greater London Authority
DEPC	Domestic Energy Performance Certificate
MHCLG	Ministry for Housing, Communities and Local Government
UN	United Nations
QGIS	Quantum Geographic Information System
GIS	Geographic Information System
VIF	Variance Inflation Factor
ANOVA	Analysis of Variance

TABLE OF CONTENTS

Introduction *1*

1. *Literature Review* *3*
    - 1.1. *Urban Green Spaces* *3*
    - 1.2. *Hedonic Regression Method* *5*
    - 1.3. *Amenities & Indicators* *9*
2. *Methodology* *12*
    - 2.1. *Research Scope and Area* *12*
    - 2.2. *Research Sample* *13*
    - 2.3. *Spatial Variables Sample* *17*
    - 2.4. *Variable Selection* *19*
    - 2.5. *Descriptive Statistics* *23*
    - 2.6. *The Hedonic Pricing Model & Hypothesis* *29*
3. *Empirical Results* *32*
4. *Limitations And Future Research* *44*
    - 4.1. *Methodological limitations* *44*
    - 4.2. *Limitations in measuring green infrastructure* *45*
    - 4.3. *Dataset limitations and omitted variables* *45*
    - 4.4. *Spatial and contextual limitations* *46*
    - 4.5. *Future research directions* *46*

Conclusion *48*

Bibliography *51*

Appendix A - London Variables Selection R Code *54*

Appendix B - Excluding Columns R Code *55*

Appendix C - Deleting Insignificant Values From Floor Area R Code *57*

Appendix D - Deleting Rows With Invalid Data On Construction Age Band R Code *58*

Appendix E - Deleting Rows Outside City Center R Code *59*

INTRODUCTION

Urbanisation has transformed the relationship between residents and the natural environment, particularly within dense metropolitan areas where access to open and green spaces has become increasingly limited. In such contexts, green infrastructure represents more than a landscape or ecological component of the urban fabric; it contributes to the quality of residential environments by providing recreational opportunities, improving environmental conditions, supporting well-being, and enhancing the overall attractiveness of neighbourhoods. However, many of these benefits are not directly reflected in conventional market transactions, making their economic value difficult to quantify. Understanding whether and how these environmental benefits are incorporated into housing markets is therefore essential for assessing the broader role of green infrastructure in sustainable urban development.

The hedonic pricing framework provides an appropriate approach for investigating this relationship. Formalised by Rosen (1974), the method considers the price of a differentiated good as the result of the combination of its individual characteristics. Applied to the housing market, this means that the value of a dwelling reflects not only its physical attributes, such as size, type, age, and quality, but also locational and environmental characteristics, including accessibility, urban services, and environmental amenities. In this context, housing prices can provide an indirect measure of the value that residents assign to characteristics that are not traded separately in the market.

Hedonic approaches have subsequently been applied to examine how environmental amenities are reflected in residential property values. Jim and Chen (2006), for example, investigated the influence of urban environmental elements on housing prices, while Panduro and Veie (2013) showed that the capitalisation of green spaces may differ according to their type, accessibility, and surrounding context. These contributions are particularly relevant to the present research, which focuses on whether proximity to significant green spaces is reflected in residential property values across different areas of London. However, the magnitude of this effect remains uncertain, as the economic importance of green infrastructure may vary considerably depending on local urban conditions, the availability of alternative amenities, neighbourhood characteristics, the type and scale of green spaces considered, and the perceived value of these amenities from the perspective of potential buyers.

Previous studies have frequently identified a positive association between proximity to parks and residential property values, suggesting that households are willing to pay a premium for improved access to green amenities. Nevertheless, this relationship is not necessarily universal. In highly centralised urban areas, property values may be primarily driven by factors such as accessibility, employment opportunities, historical character, and transport connections, reducing the relative influence of environmental amenities. Conversely, in dense residential districts where green spaces are scarce, proximity to parks may represent a significant differentiating factor in housing desirability. Therefore, the value of green infrastructure should not be considered as a fixed market effect but rather as a spatially dependent relationship

influenced by the interaction between environmental resources and the surrounding urban context.

A further methodological challenge concerns the definition of green infrastructure itself. Urban environments contain a wide range of green elements, from small landscaped areas and local gardens to large public parks. These spaces provide different levels of accessibility, recreational value, and environmental benefits, meaning that treating all green areas as equivalent may underestimate or distort their contribution to housing values. This study addresses this issue by focusing specifically on larger green spaces, using distance from parks above a defined size threshold as a proxy for access to significant green infrastructure. This approach aims to capture green areas that are more likely to function as meaningful neighbourhood amenities rather than incidental vegetation.

The objective of this research is to investigate the extent to which proximity to major green spaces influences residential property values across selected London districts. A hedonic regression model is developed using adjusted property prices per square metre as the dependent variable, while controlling for a range of structural, building-quality, and accessibility characteristics. The analysis incorporates property type, floor area, construction period, energy performance, glazing characteristics, distance from transport infrastructure, distance from educational facilities, and distance from green infrastructure. By estimating separate models for different districts, the research evaluates whether the contribution of green proximity remains consistent across urban contexts or whether its influence changes according to local market characteristics.

The results provide evidence that the relationship between green infrastructure and housing values is complex and spatially variable. While some districts demonstrate a clear economic benefit associated with living closer to major green spaces, others show weaker or different relationships, indicating that the value of greenery is influenced by broader patterns of urban development and housing market dynamics. This study therefore contributes to the understanding of environmental valuation in urban housing markets by demonstrating that the economic importance of green infrastructure depends not only on physical proximity but also on the characteristics, availability, and role of green spaces within the wider neighbourhood system.

1. LITERATURE REVIEW

Understanding the value of urban greenery is paramount for policymakers, given its myriad economic, environmental, and social benefits. At its core, the value of urban green spaces can be conceptualised as the price individuals are willing to pay for it, a dynamic perception that fluctuates over time. Unveiling the intricacies of determining the market value of urban green spaces necessitates an understanding of this fluid and subjective valuation.

The appreciation for urban greenery gained momentum in the 1970s, a period when environmental concerns became a focal point on the global agenda. The 1972 United Nations (UN) Conference on the Human Environment catalysed a shift in societal values, ushering in a new era of environmental awareness and a growing recognition of the links between urban health and green spaces. As societies became more conscious of the environmental and health implications of urban living, the value of green spaces within cities began to be seen not only as a luxury but as an essential component of urban well-being.

In parallel, rapid urbanisation over the last few decades has reshaped the dynamics of land use in cities. By 2007, more people were living in urban areas than in rural ones, a trend that has only accelerated. According to the World Bank (2023), cities now account for two-thirds of global energy consumption and more than 70% of greenhouse gas emissions, despite occupying only around 1% of the earth's land area (Hannah et al., 2018). This urban expansion has resulted in environmental pressures, such as air pollution, heat islands, and limited green space, highlighting the need for more sustainable and livable urban environments. Green spaces, in this context, are increasingly recognised as vital for mitigating these challenges and improving the quality of life for urban residents.

This chapter will delve into the existing literature on the value of urban green spaces, examining the economic, environmental, and social dimensions of this complex subject. It will explore the impact of urban greenery on property values and introduce the hedonic regression model as a method to quantify these effects. The subsequent sections will further elaborate on how this model is applied to assess the tangible benefits of urban green spaces in the housing market.

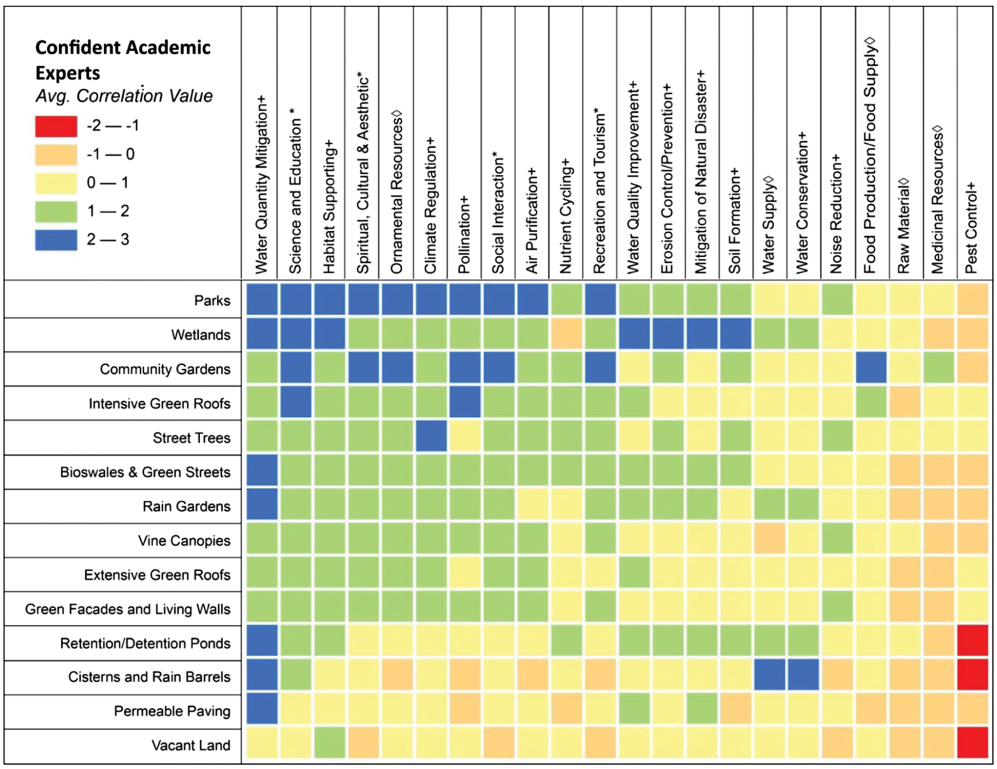
1.1. *Urban Green Spaces*

As mentioned previously, in recent decades, the quality of the urban environment has become a central issue in urban planning, directly influencing decisions on development and design. In response to the challenges posed by urbanisation, many researchers have focused on Green Infrastructure (GI), a network of natural and semi-natural features strategically integrated into urban areas. This approach offers significant environmental, social, and economic benefits, presenting a promising means of mitigating the adverse effects of urbanisation while simultaneously providing Ecosystem Services (ES) — the direct and indirect benefits people derive from ecosystems, such as flood mitigation and disaster resilience.

A substantial body of research has been dedicated to exploring the complex relationships between urban green infrastructure (GI) and ecosystem services (ES). Notably, Elliott et al. (2019) examined these relationships through the lens of public perception. Their survey employed a matrix method to investigate how respondents from diverse disciplinary backgrounds perceived the ecosystem service provision associated with different types of green infrastructure.

Rather than identifying the actual ecological relationships between GI and ES, the study explored how these relationships are understood and interpreted by citizens. The findings revealed substantial variation in perceived associations between different forms of GI and specific ecosystem services. As shown in Figure 1, larger green spaces such as parks, community gardens, and wetlands were perceived as being strongly linked to a wide range of ES. In contrast, infrastructure primarily designed for stormwater management, such as permeable paving, cisterns, and rain barrels, was perceived as providing fewer ecosystem services and in some cases showed weak or negative perceived associations. These findings highlight the diversity of public perceptions regarding green infrastructure and suggest that citizens attribute different levels of ecosystem service provision to different forms of urban greenery.

Figure 1: Urban Green Infrastructure & Ecosystem Services



Source: Elliott et al. (2019). Ambio 2020, 49:569–583 p. 575.

The importance of these findings lies not in establishing objective ecological relationships, but in understanding how citizens perceive the benefits associated with different types of green infrastructure. This perspective is particularly relevant in the context of real estate and urban development. While the ecological and social benefits of ecosystem services have been extensively documented, the value attributed to urban greenery by residents and investors is largely influenced by their awareness and perception of those benefits. Consequently, the economic

value associated with green spaces is closely linked to how individuals perceive and understand the services they provide.

A relevant contribution to the discourse on urban greenery comes from a House of Commons (2017) report on public parks in the UK. With input from over 13,000 participants across the country, the survey found that urban greenery plays a multifaceted role as a biodiversity enhancer, a mitigator of climate change, and a catalyst for community cohesion. Furthermore, the survey revealed that nearly 90% of participants believed parks had a positive impact on both individual well-being and the community. It also highlighted that over 60% of respondents typically walk to their local park, with 70% living within a 10-minute walk of one. The report also noted the significant benefits of green spaces for children's development and the health and well-being of the elderly, both directly and indirectly.

The relationship between green spaces and mental health has been extensively studied, with de Vries et al. (2003) offering notable insights. Through a linear regression analysis, they found that individuals living in greener environments reported fewer health symptoms and perceived better general and mental well-being. Remarkably, a 10% increase in local green space was associated with health benefits equivalent to reducing one's age by five years. Additionally, the introduction of new green areas was found to have a positive influence on the well-being of residents across a 3 km radius.

Taken together, these studies suggest that both the objective benefits of green infrastructure and citizens' perceptions of those benefits play an important role in shaping urban development outcomes. While ecological research establishes the ecosystem services provided by green infrastructure, public perceptions influence how these services are valued and whether citizens are willing to support related investments and policies. Understanding how perceptions translate into economic value is therefore essential, as individual preferences and choices ultimately influence the feasibility and success of urban planning initiatives. In this sense, public perception represents a key mechanism through which the benefits of green infrastructure are recognised, valued, and incorporated into urban development decisions.

1.2. Hedonic Regression Method

As green spaces continue to gain importance within urban contexts, it is essential to understand how they influence people's decisions when purchasing a home. This understanding not only aids urban planning but also provides valuable insights into how environmental amenities are perceived and valued by residents. Since green spaces are not generally traded independently in conventional markets, their economic value cannot be directly observed. The housing market, however, provides a means through which part of this value can be inferred by examining how much individuals are willing to pay for properties characterised by different levels of access to environmental amenities.

The conceptual foundation of the hedonic approach lies in the idea that consumers derive utility not simply from a good as a whole, but from the individual characteristics that the good

provides. Lancaster (1966) developed this perspective by arguing that consumer preferences are directed towards the attributes embodied in goods rather than towards the goods themselves. This idea is particularly relevant to housing, since a dwelling is a heterogeneous good composed of numerous characteristics. When choosing between properties, households simultaneously evaluate attributes such as floor area, property type, location, accessibility, neighbourhood quality, environmental conditions, and nearby amenities. Their willingness to pay for a dwelling therefore reflects the combined value they assign to these different characteristics.

Building upon this principle, Rosen (1974) formalised the hedonic pricing framework by describing differentiated goods as bundles of characteristics whose implicit values are revealed through market prices. A dwelling can therefore be represented by a vector of characteristics, as shown in (1), where z represents the complete set of attributes associated with a property, and each z_i represents an individual characteristic of the dwelling. These characteristics may include structural attributes, locational factors, accessibility measures, neighbourhood conditions, and environmental amenities.

$$z = (z_1, z_2, z_3, \ldots, z_n) \tag{1}$$

The observed market price of a dwelling can then be expressed through the hedonic price function shown in (2), where P denotes the observed property price and z_i represents the individual characteristics that jointly determine the value of the dwelling. Although buyers pay a single market price for a property, the hedonic approach seeks to identify the implicit contribution of each characteristic to the overall price.

$$P = P(z) = P(z_1, z_2, \ldots, z_n) \tag{2}$$

The relationship between an individual characteristic and the property price can be represented through the marginal implicit price. For a specific attribute z_k, this is defined as the partial derivative of the hedonic price function with respect to that characteristic, as shown in (3), where pk(z) represents the marginal implicit price of characteristic zk. This value indicates the change in property price associated with a marginal change in a specific attribute, while all other characteristics remain constant.

$$p_k(z) = \frac{\partial P(z)}{\partial z_k} \tag{3}$$

Under the equilibrium assumptions of the hedonic framework, this marginal implicit price corresponds to the household's marginal willingness to pay for that characteristic (Rosen, 1974). In practical terms, if properties located closer to a park systematically command higher prices than otherwise comparable properties located farther away, the difference in observed prices may be interpreted as the implicit value associated with access to that environmental amenity. The hedonic method therefore provides a mechanism for estimating the economic value of characteristics, such as environmental quality or proximity to green spaces, that do not have an independently observed market price.

Analysing house prices through this framework remains complex because dwellings differ simultaneously across numerous dimensions. A reliable hedonic model must therefore account for the principal characteristics that may influence market value. Structural characteristics such as floor area, property type, age, and building quality generally interact with locational and neighbourhood attributes, including accessibility to employment, transport infrastructure, schools, environmental quality, and public amenities. As described by O'Sullivan (2012), proximity to employment opportunities or high-performing schools may generate price premiums, whereas characteristics such as higher crime rates, toxic waste sites, or exposure to noisy highways may produce price discounts. Hedonic regression attempts to distinguish these effects statistically by estimating the contribution of each characteristic while controlling for the influence of the others.

The hedonic pricing method belongs to the broader family of revealed-preference approaches. Rather than asking individuals directly how much they would be willing to pay for a particular amenity, revealed-preference methods infer preferences from actual market behaviour. In the housing context, observed transactions therefore provide information about how households trade off different characteristics when selecting a dwelling. This differs from stated-preference approaches, which rely on surveys or hypothetical scenarios in which individuals are directly asked how much they would be willing to pay for a particular improvement or attribute. While both approaches can provide useful information on preferences, the hedonic method is specifically based on the choices and prices observed in real markets.

This characteristic-based approach has been widely applied to real-estate markets, particularly to investigate characteristics whose economic value is not directly observable. Jim and Chen (2006), for example, applied a hedonic pricing model to residential properties in Guangzhou, China, examining environmental elements including green-space views, proximity to wooded areas and water bodies, and exposure to traffic noise. Their findings demonstrated that environmental characteristics can be reflected in residential property prices, illustrating how market transactions can be used to estimate the implicit value of urban environmental amenities. Similarly, Fávero (2011) examined the relationship between urban amenities and dwelling prices in São Paulo through a multilevel modelling approach, highlighting the importance of considering both property characteristics and spatial differences between districts.

The application of hedonic models has increasingly benefited from the development of Geographic Information Systems (GIS), which allow spatial characteristics to be measured with greater precision. GIS does not represent a separate valuation method but rather provides tools for constructing variables describing the spatial relationship between properties and surrounding amenities. Distances to parks, transport infrastructure, educational facilities, water bodies, and other neighbourhood characteristics can consequently be incorporated directly into the hedonic price function. Osland et al. (2022), for instance, combined housing information with GIS-derived environmental variables to analyse property prices in Oslo. Their research demonstrated that different measures of access to and the presence of environmental amenities can contribute differently to housing values, illustrating the importance of carefully defining spatial variables within hedonic models.

The use of hedonic regression is therefore particularly relevant for the valuation of green infrastructure. Environmental amenities such as parks, urban forests, and water bodies generate benefits that are not normally purchased separately by residents but may nevertheless influence their residential choices. If households value these amenities, part of that value may be capitalised into surrounding property prices. Studies such as Jim and Chen (2006) and Panduro and Veie (2013) demonstrate how hedonic models can be used to identify this relationship. In particular, Panduro and Veie (2013) show that green space should not necessarily be treated as a homogeneous amenity, as accessibility, maintenance, type, and surrounding land uses can generate different implicit values. The hedonic approach therefore provides an appropriate framework for the present research, which seeks to examine whether proximity to significant green spaces is reflected in residential property values in London.

Despite its widespread application, the hedonic approach is subject to limitations that can be grouped into at least two broad categories. The first relates to the conceptual foundations of the method. Hedonic models are based on the assumption that the value of a property can be decomposed into the implicit values of its individual characteristics, such as location, structural features, environmental quality, and access to amenities. While this framework has proven highly influential in the analysis of housing markets, consumers may perceive and evaluate a dwelling as an integrated good rather than as a simple aggregation of separable attributes. Consequently, the utility derived from a property may depend not only on the individual characteristics themselves but also on interactions between them, making it difficult to isolate the precise contribution of each attribute to overall market value.

The second category of limitations concerns the econometric implementation of hedonic models. Accurate estimation requires the researcher to control for the relevant determinants of housing prices and to satisfy a range of statistical assumptions regarding model specification and error terms. If relevant characteristics are omitted and are correlated with variables included in the model, coefficient estimates may be biased. This issue is particularly important in housing markets, where many characteristics related to property quality, neighbourhood conditions, or household preferences may be difficult to observe.

These limitations became clearer in later studies by Earnhart (2002) and Kuminoff et al. (2014), which contributed to a more nuanced understanding of the complexities associated with hedonic analysis. Earnhart (2002) argues that, although revealed-preference approaches are grounded in observed behaviour, their results depend strongly on adequately representing the factors influencing housing choices. The study also highlights difficulties associated with defining the set of housing alternatives available to households and proposes combining revealed and stated information to improve the analysis of residential decisions. Similarly, Kuminoff et al. (2014) draw attention to the interpretation of capitalisation effects and to the difficulties involved in using observed changes in property values as direct measures of willingness to pay, particularly when markets adjust following changes in public goods or environmental characteristics.

These considerations do not undermine the usefulness of hedonic regression but highlight the importance of interpreting its results carefully. The method remains particularly valuable for

examining how structural, locational, and environmental characteristics are reflected in observed property prices. When combined with appropriate control variables and spatial information, hedonic analysis provides a useful framework for investigating the implicit economic value associated with urban amenities and for understanding how such amenities contribute to housing market outcomes.

1.3. Amenities & Indicators

Hedonic regression is also a valuable tool in estimating the price of individual house features whose impact cannot be directly observed, such as the number and characteristics of bathrooms, bedrooms, and other amenities. It achieves this by applying multiple regression analysis to a pooled sample of many dwellings. For example, a study conducted by Sirmans et al. (2005) shows that hedonic pricing models incorporate various variables related to house features, location, marketing, and finance. The study proposes eight categories: (i) structural features (e.g. lot size, square footage, age); (ii) internal features (e.g. baths, fireplace, air-conditioning); (iii) external features (e.g. garage spaces, deck, pool); (iv) natural environmental features (e.g. lake view, lakefront, ocean view); (v) neighbourhood and location (e.g. location, crime rates, distance); (vi) public services (e.g. school district, percentage of school district minorities, public sewer); (vii) marketing, occupancy, and selling factors (e.g. assessed condition, vacant, time-on-the-market); and (viii) financing issues. The research concludes that most characteristics tend to be directly proportional to selling prices, except for crime rates, age, percentage of school district minorities, and vacancy.

Sirmans et al. (2005) classification, however, is not without its limitations, as the literature presents several variations in classification schemes and indicators used. According to a comprehensive review by Schaeffer et al. (2018), who analysed 328 articles published over 40 years, most studies use simple indicators reflecting unidimensional natural attributes. More complex indicators, such as multidimensional measures (which focus on proximity to specific attributes such as lakes or parks), ecological indicators (derived from ecological science, like landscape metrics), or subjective views of the surrounding natural environment, are less frequently employed. The authors concluded that considering a broader range of indicators allows researchers to gain a deeper understanding of the multifaceted influences of natural and environmental amenities on various outcomes.

It is important to note that the impact of green spaces on housing prices can vary significantly across different geographical and cultural contexts. For instance, while parks and lakes may yield a price premium in cities like Aalborg (Panduro and Veie, 2013) and Guangzhou (Jim et al., 2006), similar green spaces in other urban environments may not exhibit the same influence. This variation can be attributed to cultural differences in how green spaces are perceived and utilised. In some cities, green spaces may be seen primarily as leisure areas, while in others, they may hold additional symbolic value related to environmental sustainability or community cohesion. Moreover, the accessibility and functionality of these spaces can differ substantially depending on local urban planning practices, the density of development, and the availability of other

amenities. For example, in densely populated cities where space is scarce, green spaces may have a more pronounced effect on housing prices, as they provide valuable respite from the built environment, enhancing both physical and mental well-being. In contrast, in less densely populated areas, the impact of green space may be diluted, as residents may not place as high a value on proximity to these spaces due to alternative recreational options or more abundant outdoor spaces.

The categorisation of green spaces is crucial for understanding their varying impact on property prices. Studies such as those by Panduro and Veie (2013) and Fan et al. (2015) highlight the importance of distinguishing between different types of green spaces based on their size, function, and accessibility. Larger parks and lakes, for instance, tend to be more valued in residential areas, especially in urban environments where open space is limited. However, the presence of smaller, more localised green spaces, such as community gardens and common areas, can have a more significant impact in specific contexts, particularly in areas where residents value immediate access to recreational spaces. In this sense, the concept of 'green buffers,' such as those surrounding industrial areas, which often have limited access and poor maintenance, illustrates how the presence of green space alone does not guarantee positive effects on property values. On the contrary, these spaces can be perceived as undesirable, leading to a negative correlation with house prices. This complexity underscores the necessity of considering both the quantity and quality of green space in hedonic regression models, as they shape the residential experience in different ways.

The authors define accessibility to green spaces using a series of variables specific to each type. For example, proximity is measured using Euclidean distance, with 100m intervals from properties to various green spaces. Exceptions are made for 'common areas', where size is more important than proximity due to their closeness to residential areas, and lakes, which in Aalborg are less accessible for apartments, so they are omitted from apartment models. A high cut-off distance of 600 metres is set for parks and natural areas deemed attractive, reflecting the practical walking limit for residents. For club goods and smaller green spaces, a lower cut-off of 300 metres is used.

Panduro and Veie's analysis found that parks and lakes are associated with higher housing prices, while green buffers, typically located around industrial areas, have negative effects due to limited access, poor maintenance, and undesirable neighbouring properties. The study also highlighted differences in the capitalisation of green space between apartments and houses. For example, common areas were found to have a significant positive effect on apartment prices, but no such effect was found for houses, suggesting that private gardens in houses diminish the importance of shared spaces.

This distinction in the impact of green spaces on different property types aligns with findings in other studies. For instance, Fan et al. (2015) categorised green areas in two adjacent Central California cities into four types: Regional Parks, Seasonal Parks, Community Parks, and Golf Courses, based on distance and size. Kong et al. (2007) used a similar classification for Jinan City (China), categorising spaces into Parks, Plazas, and Scenery Forests, with a focus on accessibility,

size, and quality. Kim et al. (2019) classified green spaces into Living Zone Parks (subdivided into neighbourhood parks, children's parks, and small parks) and Theme Parks (subdivided into historic, cultural, waterside, and sports parks). These studies illustrate the varying aspects and levels of amenities provided by different categories of green space.

An often-overlooked factor in determining the value of green spaces is public perception. The extent to which environmental amenities are capitalised into housing prices depends not only on their objective characteristics but also on how these characteristics are perceived and valued by households. Benefits such as improved environmental quality, recreational opportunities, aesthetic enjoyment, or climate regulation may influence housing demand only to the extent that potential buyers recognise and appreciate them. Consequently, differences in public awareness and preferences can lead to variation in the willingness to pay for similar green amenities across locations and populations. This perspective is consistent with the findings of Panduro and Veie (2013), who argue that the value attributed to urban green spaces depends not only on their physical characteristics but also on how these characteristics are perceived by residents and reflected in housing market behaviour.

Sociodemographic factors, as identified in studies like that of Su et al. (2021), play a pivotal role in shaping how different groups within a city value green spaces. In megacities such as Beijing and Shanghai, the priorities of migrant workers often centre around economic factors like job availability and transportation access, while the 'creative class' tends to place greater emphasis on the quality of their living environment, particularly the availability of natural amenities. This socioeconomic divide highlights the complex relationship between housing prices and green space. The value attributed to environmental amenities may vary systematically across socioeconomic groups, suggesting that the effect of green space is not necessarily independent of household characteristics such as income, occupation, education, or migration status. Consequently, treating the housing market as a homogeneous entity may obscure important sources of variation in willingness to pay. Understanding these differences is therefore essential for identifying submarket-specific preferences and obtaining more accurate estimates of the value associated with green amenities.

In conclusion, while it is impossible to account for every characteristic affecting the price of a dwelling, particularly given the complexity of housing markets and the possibility of omitted variables, understanding the heterogeneity of amenities and their effects on surrounding areas is crucial. The challenge lies not in the existence of individual differences in preferences, but in ensuring that the most relevant determinants of housing prices are adequately represented within the model. Recognising the diverse characteristics and functions of green spaces allows hedonic analyses to provide more accurate estimates of their contribution to property values.

2. METHODOLOGY

2.1. Research Scope and Area

This research investigates the multifaceted relationship between urban greenery and its broader socio-economic value, employing hedonic regression analysis to examine how specific attributes —such as aesthetic quality, accessibility, and environmental sustainability—influence both property values and residents' well-being. Rather than treating green spaces as uniform urban amenities, the study approaches them as differentiated assets whose perceived and measurable qualities generate distinct economic and social effects. By exploring how individuals experience, evaluate, and ultimately assign value to urban green environments, this research contributes to ongoing debates in urban economics, environmental planning, and public policy. Its broader aim is to support the development of accessible, resilient, and high-quality green infrastructure capable of meeting environmental objectives while responding to the diverse expectations of urban communities.

The choice of London as the case study is grounded in a combination of historical, analytical, and practical considerations that make it particularly well suited to this investigation. As one of the world's leading global cities, London presents a uniquely complex urban landscape in which green spaces have long played a defining role in shaping patterns of urban life. From medieval churchyards and royal parks to Victorian garden squares and contemporary ecological corridors, the city's green infrastructure reflects centuries of evolving relationships between nature and urban development. This historical continuity provides an exceptional framework through which to examine how urban greenery has adapted to shifting social priorities, demographic pressures, and planning paradigms over time.

At the same time, London's highly dynamic and internationally significant property market offers fertile ground for analysing the economic implications of environmental amenities. Housing values in the city are shaped by a dense interplay of locational, structural, and neighbourhood characteristics, making London an especially relevant context for hedonic modelling. Within such a competitive and stratified market, the contribution of green spaces to residential desirability becomes both measurable and policy-relevant. Investigating these dynamics in London allows for a nuanced understanding of how environmental quality is capitalised into real estate values, while also revealing how such benefits may be unevenly distributed across socio-economic groups and geographic areas.

Another decisive factor in the selection of London lies in its exceptional data availability. The city is among the most extensively documented urban environments in the world in terms of planning, housing, and demographic information. Publicly accessible datasets from institutions such as the Greater London Authority, the London Datastore, and the Office for National Statistics provide a robust empirical basis for quantitative analysis. These sources include detailed property transaction records, comprehensive inventories of green spaces, socio-economic indicators, and resident survey data. The breadth and transparency of such

information strengthen the methodological reliability of the research and allow for a granular investigation of how urban greenery interacts with other determinants of housing values within London's diverse urban fabric.

Beyond its empirical suitability, London also represents a strategically important case in light of the wider challenges confronting contemporary cities. The city contains a large and diverse network of green spaces, ranging from major royal parks and metropolitan open spaces to smaller neighbourhood parks and ecological corridors. However, these amenities are not evenly distributed across the urban landscape. Significant variations exist in terms of accessibility, size, quality, and proximity to residential areas, creating a spatially differentiated environment that is particularly well suited for assessing the value of green infrastructure. This heterogeneity provides the variation necessary to examine how different forms of urban greenery are capitalised into housing prices and experienced by residents across distinct neighbourhood contexts.

Furthermore, London's social, economic, and demographic diversity offers an opportunity to investigate how preferences for green amenities may vary across different population groups and neighbourhoods. Given the city's highly segmented housing market, residents are likely to differ in the way they experience, evaluate, and assign value to urban green spaces. This variation is particularly relevant for a study concerned with perceived benefits and willingness to pay, as it allows for the exploration of how environmental amenities are reflected in housing prices across different spatial and socioeconomic settings. Examining these dynamics within London therefore provides valuable insights into the relationship between urban greenery, residential preferences, and property values.

Given its global prominence, London functions not merely as a localised case study, but as a reference point for wider international debates on sustainable urban development. Insights derived from its experience can inform strategies in other metropolitan regions facing similar tensions between land scarcity, housing demand, and ecological imperatives. Understanding how green spaces contribute to quality of life, economic value, and urban resilience in London therefore offers lessons that extend well beyond the city itself, reinforcing the relevance of this research for planners, policymakers, and scholars concerned with the future of cities worldwide.

2.2. Research Sample

In order to examine how urban green spaces affect London's housing market, this research employs a dataset made available by Chi et al. (2021) at University College London (UCL). This includes transaction data for England and Wales, featuring the Land Registry's Price Paid Data from 1995 to 2021, published by the Greater London Authority (GLA) at London Datastore. Additionally, it incorporates Domestic Energy Performance Certificate data published by the Ministry for Housing, Communities and Local Government (MHCLG), along with three variables created by the first author. In total, the dataset records almost 18 million transactions and 104 variables. Due to the vast size of the original CSV file, the dataset variables were filtered

using Anaconda Spyder software and Python language to tailor the data for the purposes of this research.

The dataset is highly comprehensive and contains a wide range of variables relating to housing transactions, energy performance, location, and neighbourhood characteristics. However, as it is primarily based on administrative records of market transactions, it provides limited information on certain internal property attributes, such as the quality of finishes, renovation status, architectural design, or detailed interior layouts. The absence of these variables does not necessarily undermine the analysis, as omitted characteristics that are independently distributed can be absorbed by the error term without biasing coefficient estimates. Nevertheless, potential concerns may arise if unobserved property characteristics are systematically correlated with environmental amenities or neighbourhood attributes included in the model. For example, if higher-quality dwellings are disproportionately concentrated near attractive green spaces, part of the estimated effect of urban greenery could reflect these unobserved differences. Despite this consideration, the breadth of the dataset and the inclusion of numerous structural and locational controls make it well suited for investigating the relationship between urban green spaces and housing prices.

The first stage of the data preparation process involved developing a dedicated code to filter the dataset so that only variables relating to the city of London were retained (see Appendix A - London Variables). This initial operation was essential in narrowing the analytical scope and ensuring geographic consistency across the observations. Given the breadth of the original dataset, a second script was subsequently implemented to remove variables deemed irrelevant to the objectives of this study—such as cost efficiency potential indicators and other attributes unrelated to housing valuation—thereby reducing the overall size and complexity of the CSV file (see Appendix B - Excluding Columns).

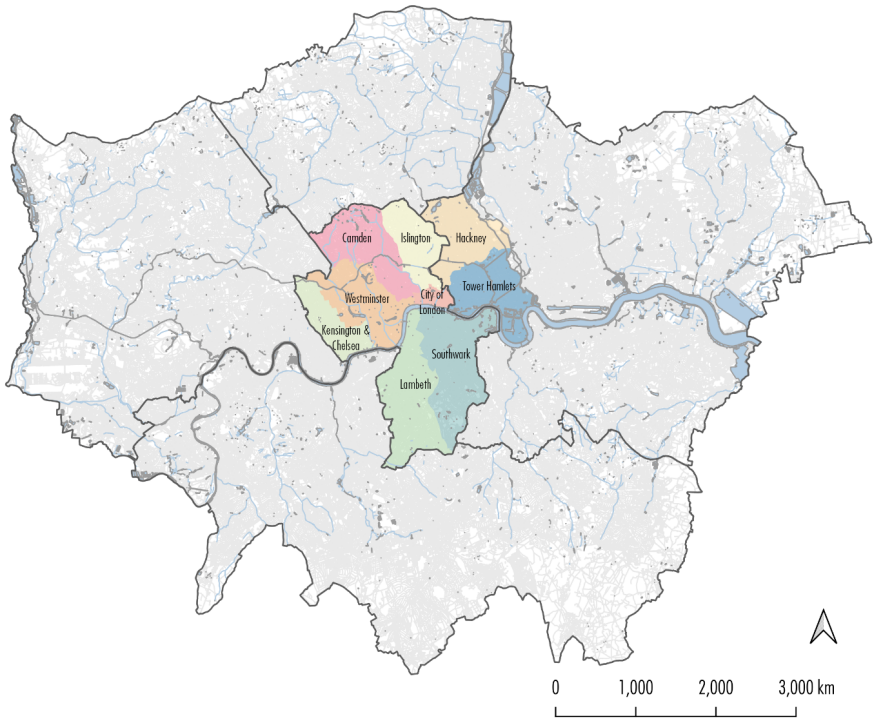
Following this preliminary filtering, further data-cleaning procedures were undertaken to improve the reliability of the sample. Rows containing zero values or otherwise invalid entries were systematically deleted, particularly in relation to floor area and construction age band variables (see Appendix C - Deleting Insignificant Values From Floor Area and Appendix D - Deleting Rows With Invalid Data On Construction Age Band). These steps were necessary to prevent distortions in the econometric analysis, as anomalous or incomplete records can introduce bias and weaken the explanatory power of the model.

A key methodological concern at this stage was maintaining data quality. In empirical research, especially when relying on large-scale secondary datasets, inconsistencies, omissions, and outliers can significantly affect the robustness of results. For this reason, considerable attention was devoted to ensuring that the dataset was as accurate and internally coherent as possible. By excluding invalid observations and refining the sample, the study reduces the likelihood of spurious relationships emerging between urban greenery and property values. This process strengthens the credibility of the subsequent hedonic regression analysis, allowing the findings to rest on a more dependable empirical foundation.

The final step in the preparation process consisted of spatially restricting the analysis to the city centre, with the aim of mitigating potential interferences arising from negative externalities associated with living at greater distances from London's Business Centre, as discussed in the previous chapter (O'Sullivan, 2012) (see Appendix E - Deleting Rows Outside City Center). This geographical delimitation is methodologically significant, as it concentrates the analysis on properties for which proximity to urban amenities—including green spaces—is more likely to constitute a meaningful determinant of value. In other words, it reduces the noise introduced by peripheral housing markets, where commuting constraints, transport dependency, and suburban characteristics may overshadow the influence of environmental amenities.

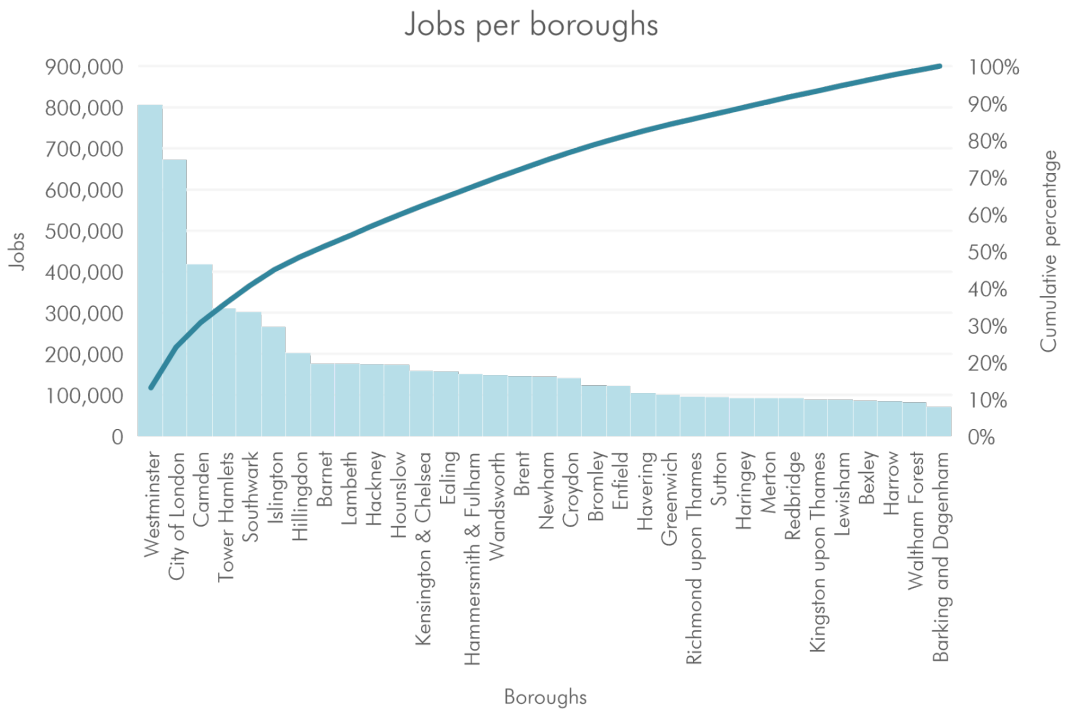
As a result, the research focuses on districts in Central London and selected parts of East London, as illustrated in Figure 2. The inclusion of eastern areas was not incidental, but rather a deliberate analytical choice supported by data from the Office for National Statistics (2023), which highlights the distribution of available jobs across the city's regions (Figure 3). East London has undergone substantial transformation in recent decades, evolving into an increasingly important residential and commercial zone. Urban regeneration projects, infrastructural investment, and demographic change have redefined its role within the metropolitan economy. In this context, access to green spaces and high-quality public infrastructure has become a critical factor in shaping patterns of residential desirability and, by extension, property values.

Figure 2: London division and research focus area



Source: made by author using GIS software. Contains National Statistics data © Crown copyright and database right [2015]

Figure 3: Jobs per borough



Source: made by author using data available at London Datastore issued by ONS, 2022.

Conversely, the exclusion of certain boroughs—such as Hillingdon and Barnet—formed an equally important component of the methodological design. Although these boroughs may display relatively high employment figures, their greater distance from central London and their broader territorial extent introduce additional spatial and socio-economic variables that could confound the analysis. Their inclusion would risk weakening the internal comparability of the sample by incorporating housing market dynamics that are more characteristic of outer suburban areas than of the central metropolitan core.

By excluding them, the study is better positioned to isolate the effects of urban greenery within a more homogeneous and representative segment of London's housing market. This targeted approach enhances analytical precision and ensures that the estimated relationship between green spaces and property values is examined in urban contexts where accessibility, density, and amenity proximity are central determinants of residential pricing. In doing so, the research preserves its focus on those areas where the valuation of urban greenery is most visible and economically relevant.

Once the dataset had been reduced to a manageable and methodologically coherent size, additional variables were incorporated in order to enrich its analytical depth. This was achieved through the XLOOKUP tool in Excel, which enabled the integration of supplementary geographical information by matching each property's postcode with a secondary dataset provided by the Office for National Statistics (2011). The additional file included key locational variables such as latitude, longitude, and borough name, thereby embedding each observation within a clearer spatial framework.

The inclusion of these geographical indicators represents a crucial enhancement for the subsequent stages of analysis. In hedonic modelling, locational precision is fundamental, as property values are strongly influenced by spatial characteristics and neighbourhood context. By assigning geographic coordinates and borough identifiers to each record, the study gains the capacity to examine not only the direct relationship between green spaces and housing prices, but also the spatial distribution of these effects across different urban environments. Such variables are particularly valuable when exploring proximity-based measures, accessibility patterns, and borough-level disparities in environmental amenities.

Moreover, the integration of these data points strengthens the interdisciplinary character of the research by linking economic analysis with urban geography. It enables a more nuanced understanding of how environmental and locational factors interact within London's housing market, moving beyond purely structural determinants of value. In practical terms, this enriched dataset creates the empirical basis for more refined statistical modelling and spatial interpretation.

Finally, it is important to underline that London is internationally recognised for its high degree of data transparency. A wide range of publicly accessible, royalty-free datasets is made available for scientific and policy-oriented research, greatly facilitating empirical investigation. This openness not only supports the methodological rigour of the present study, but also enhances its replicability. The ability for other researchers to access equivalent datasets ensures that findings can be verified, challenged, or extended in future studies. Such transparency is especially significant in urban research, where reproducibility remains essential for building cumulative and reliable knowledge. In this sense, London's data ecosystem constitutes not merely a practical advantage, but also a foundational element in the scientific credibility of the research.

2.3. Spatial Variables Sample

To enhance the analytical usability of the dataset, the integration of geo-referenced information was identified as an essential step, in line with the methodological considerations discussed in the previous chapter (see Osland et al., 2022). In its original form, the dataset lacked sufficiently detailed location-based variables, limiting its capacity to capture the spatial determinants of housing value. This absence is particularly significant in the context of urban property markets, where proximity to amenities such as parks, transport infrastructure, and educational institutions often exerts a measurable influence on residential prices. To address this limitation, a comprehensive geo-spatial enrichment strategy was implemented, combining the use of QGIS (Quantum Geographic Information System) software with additional data sourced from authoritative and open-access repositories.

Two principal sources were employed to integrate geo-referenced information into the dataset. The first consisted of data obtained from the London Datastore platform, issued by the Greater London Authority (2015) and the Census Information Scheme (2015). These datasets included key shapefiles such as Statistical Boundaries, Allotments, and Census information, all of which are

fundamental for mapping and analysing urban patterns across the metropolitan area. The incorporation of these spatial layers provided a detailed territorial structure through which demographic, environmental, and infrastructural variables could be interpreted more effectively. In practical terms, these data enhanced the capacity to examine how the geographic distribution of amenities and neighbourhood characteristics contributes to housing market outcomes.

The second source involved the use of data from OpenStreetMap (2023), specifically in Layered GIS Format. This dataset included roads, railways, and points of interest, thereby complementing the official datasets with a broader and more dynamic representation of urban infrastructure. OpenStreetMap was selected due to its open-source character, extensive coverage, and continuous updating process, which collectively make it one of the most versatile geographic data resources available for urban analysis. Its inclusion ensured access to detailed locational features that are often unavailable through proprietary or institutionally limited sources, thereby strengthening the comprehensiveness of the spatial framework.

Within this analytical process, QGIS served as the central platform for integrating these diverse datasets with the pre-existing Excel database containing real estate transaction records and price-paid information. Through the application of the "Distance to Nearest Hub" tool, which calculates the shortest path from a given point to a selected hub of interest, it became possible to attach proximity-based indicators to each property observation. This procedure transformed the original dataset into a geo-referenced database, thereby introducing a spatial dimension that is indispensable for understanding urban housing dynamics.

The resulting geo-referenced dataset includes a set of critical spatial variables that substantially expand the explanatory potential of the research. Among these, three indicators were of particular importance:

- **Dist_Metro**: the distance from each dwelling to the nearest metro station;
- **Dist_Educational**: the distance from each dwelling to the nearest educational institution, including schools, colleges, and universities;
- **Dist_Greenery**: the distance from each dwelling to the nearest green space, such as parks or public gardens.

These variables provide a structured basis for examining how access to essential urban infrastructure affects residential property values. In the context of hedonic regression, such measures are particularly valuable because they translate locational advantages into quantifiable metrics that can be statistically tested. Proximity to transport nodes may reflect accessibility and commuting efficiency; closeness to educational institutions may signal family-oriented desirability or long-term neighbourhood stability; and access to green spaces captures both environmental and recreational value.

More broadly, the incorporation of these indicators allows the analysis to move beyond conventional structural housing attributes and engage with the spatial logic of urban life. It enables a more nuanced interpretation of how environmental and infrastructural amenities are capitalised into property prices, thereby capturing the complex geography of value formation

within London's housing market. This geo-referenced approach is therefore not simply a technical enhancement, but a methodological necessity for understanding the interplay between space, accessibility, and urban quality of life.

2.4. Variable Selection

Building upon the studies examined in the previous chapter, the selection of variables for this research is grounded in the classification of housing characteristics proposed by Sirmans et al. (2015). This framework is widely recognised for its methodological robustness in identifying the principal structural, aesthetic, and locational determinants of property values. Its relevance lies in its capacity to organise housing attributes into coherent analytical dimensions, thereby supporting a systematic interpretation of market behaviour. In the context of this study, it provides the conceptual foundation for structuring the explanatory variables used in the hedonic regression model.

At the same time, the available dataset imposes important limitations on the scope of analysis. Since it is primarily derived from private market transactions, detailed internal characteristics of properties—such as maintenance condition, spatial layout, refurbishment quality, and finer domestic features like the number of bathrooms—are not comprehensively represented. These omissions restrict the ability to capture the full spectrum of factors that may influence buyer preferences and, consequently, market valuation. As a result, the research places greater emphasis on observable structural and locational characteristics, which remain consistently available and quantifiable across the dataset.

The selected variables are organised into four principal categories, each offering a distinct lens through which to assess the relationship between urban greenery and housing prices.

Structural features represent the most immediate and measurable determinants of market value. Among these, **Area** is one of the most significant predictors of price, as larger properties generally provide greater living capacity, storage potential, and long-term adaptability. In urban contexts such as London, where land scarcity intensifies competition for residential space, floor area often carries substantial economic weight. Larger homes are therefore more likely to command higher prices, not only because of their functional advantages, but also because they offer greater flexibility for households and potential for future modification.

Closely related is **Property Type**, which captures differences in residential form and market segmentation. Detached houses, semi-detached properties, terraced houses, and flats each appeal to different buyer groups and embody distinct lifestyle preferences. Detached homes, for example, are often associated with privacy, exclusivity, and family-oriented living, whereas flats may attract younger professionals, first-time buyers, or investors due to their comparatively lower entry costs and reduced maintenance demands. In a heterogeneous housing market such as London's, these distinctions are particularly relevant, as they reflect broader socio-economic and demographic patterns.

A further structural indicator is **Construction Age Band**, which provides insight into a property's architectural identity, maintenance requirements, and perceived desirability. Older properties may benefit from historical character and distinctive design features, qualities that can attract buyers seeking heritage value. However, they may also involve higher renovation costs and lower energy efficiency. Conversely, newer developments tend to be associated with modern standards, technological integration, and reduced upkeep, making them especially attractive to buyers prioritising convenience and sustainability. Age therefore functions not merely as a chronological measure, but as a proxy for multiple dimensions of housing quality.

Similarly, **Energy Rating** has become an increasingly relevant determinant of residential value, particularly in the context of rising environmental awareness and energy costs. Properties with higher energy efficiency ratings are often perceived as more sustainable, cost-effective, and aligned with contemporary regulatory expectations. In London, where utility costs are substantial and sustainability credentials are increasingly marketable, energy performance can significantly shape buyer decisions. It also reflects a broader shift in housing markets toward the valuation of long-term operational efficiency alongside traditional structural attributes.

Internal features, though more limited in the dataset, are represented through **Glazed Areas**, which capture the proportion of windows and transparent surfaces within the dwelling. This variable serves as an indirect measure of natural light availability and openness, both of which are highly valued in dense urban settings. Greater exposure to daylight contributes to perceived spaciousness, psychological well-being, and aesthetic appeal. In metropolitan contexts where access to sunlight may be constrained by surrounding built density, properties with extensive glazing often enjoy a competitive advantage. This is particularly true where such features also provide visual connections to urban landscapes or nearby greenery.

External features reflect the property's relationship with essential urban infrastructure. **Proximity to Metro Stations** is especially relevant in London, where public transport is a central component of everyday mobility. Reduced commuting time, improved accessibility, and integration into the wider metropolitan network all contribute to residential desirability. Properties located near major transport lines typically command price premiums, as accessibility remains one of the strongest predictors of urban housing demand.

Likewise, **Proximity to Educational Institutions** captures the value associated with access to schools, colleges, and universities. For families, such proximity often translates into educational opportunity and convenience; for students and academic staff, it can reduce daily travel burdens. In areas with concentrations of highly regarded institutions, this variable can substantially influence local demand and property pricing dynamics.

The fourth category, natural environmental features, forms the conceptual core of this research. Following the multidimensional distance indicator proposed by Schaeffer et al. (2018), the study incorporates **Proximity to Green Areas** as a key explanatory variable. Urban greenery is understood not merely as a recreational amenity, but as an environmental asset capable of enhancing quality of life, supporting health, and improving neighbourhood attractiveness. Parks,

gardens, and other green spaces contribute to social interaction, psychological restoration, and ecological resilience. Consequently, properties situated near high-quality green areas often command higher values, reflecting the premium placed on environmental accessibility within urban markets.

This variable is especially significant because it bridges economic valuation with broader urban planning concerns. Unlike traditional housing characteristics, the presence of greenery represents a collective environmental good whose benefits extend beyond individual ownership. Its inclusion in the model therefore allows the analysis to capture both private and public dimensions of residential value.

Furthermore, the influence of green areas on property values may depend not only on proximity itself, but also on qualitative characteristics such as maintenance standards, accessibility, safety, biodiversity, recreational functions, and aesthetic appeal. As suggested by Panduro and Veie (2013), the economic value associated with green spaces may vary considerably according to their quality and perceived utility. Similarly, studies by Kong et al. (2007) and Fan et al. (2015) emphasise that the size and typology of green spaces can significantly affect their impact on surrounding property prices. Large, well-maintained urban parks may generate stronger positive externalities than smaller or poorly maintained green areas, thereby producing differentiated effects that may not be fully captured through a simple distance-based variable such as Dist_Greenery.

To address this limitation and more accurately capture the impact of larger green spaces, the variable was refined by restricting the analysis to parks exceeding 40,000 square metres. This adjustment ensures that only substantial green areas—those more likely to generate meaningful environmental, social, and recreational benefits—are included in the analysis. In densely urbanised contexts such as London, larger parks are expected to exert a more pronounced influence on residential well-being and, consequently, on property values. By narrowing the definition of green spaces in this way, the revised specification aims to produce a more precise estimation of the relationship between proximity to significant green infrastructure and housing prices, thereby offering a more direct test of the underlying hypothesis.

Despite the robustness of the selected variables, several limitations must be acknowledged. The absence of detailed internal property attributes constrains the explanatory depth of the model, particularly in cases where interior quality strongly influences market perceptions. This may result in an incomplete representation of buyer decision-making processes, especially in high-end or highly differentiated market segments.

A further limitation lies in the operationalisation of accessibility through distance alone. While proximity to metro stations, schools, and green spaces provides a quantifiable measure of locational advantage, it does not account for differences in quality, maintenance, reputation, or actual usability. A small, poorly maintained park is unlikely to generate the same value premium as a large, well-designed urban green space, yet both may appear equivalent in a purely distance-

based framework. This simplification risks understating the nuanced ways in which environmental amenities affect housing values.

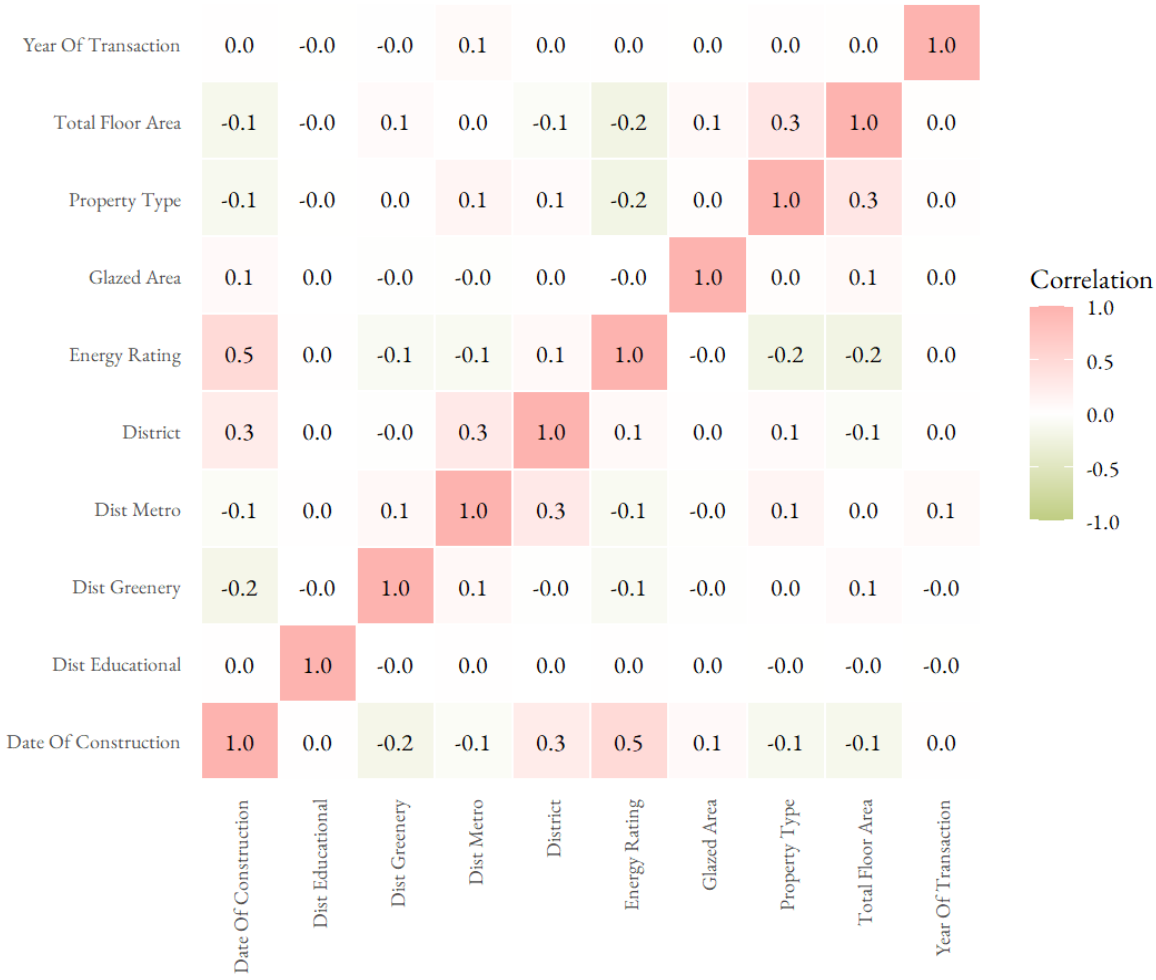
Additionally, the concentration on central and highly accessible urban areas—while methodologically justified—limits the generalisability of findings to peripheral or suburban contexts. In outer boroughs, the role of green space may differ substantially, as larger private gardens, lower density, and alternative commuting patterns reshape the relative importance of public greenery.

Nevertheless, the incorporation of spatial variables such as proximity to transport, education, and green infrastructure provides a dynamic and context-sensitive framework for understanding property valuation. These measures strengthen the model by embedding housing transactions within the lived geography of the city, thereby moving beyond static structural characteristics. Ultimately, the chosen variables ensure that the analysis remains firmly anchored in the realities of London's urban housing market, while offering a rigorous basis for assessing the economic significance of urban greenery.

Despite the robustness of the selected variables, it is useful to empirically examine their interrelationships to assess potential redundancy and confirm their distinct contributions. The correlation matrix below indicates generally low correlations among the variables, with most coefficients falling between -0.2 and 0.3, suggesting that each measure captures a unique dimension of housing value, and confirming that they provide explanatory power. The exception is a moderate positive correlation ($r = 0.5$) between Energy Rating and Date of Construction, which reflects the expected pattern that older properties tend to exhibit lower energy efficiency, whereas newer constructions often meet contemporary energy standards. Both variables remain conceptually and analytically relevant: construction age conveys structural, architectural, and maintenance-related characteristics, while energy rating captures operational efficiency and sustainability considerations that can independently influence buyer preferences.

Taken together, the correlation analysis empirically validates the multidimensionality of the selected variables and supports their inclusion in the model as shown in Figure 4. Structural, locational, and environmental characteristics are therefore represented in a manner that captures both their unique and complementary contributions to property valuation, ensuring that the hedonic regression can meaningfully assess the economic significance of urban greenery within London's complex housing market.

Figure 4: Correlation Matrix



Source: made by author.

2.5. *Descriptive Statistics*

To identify potential outliers and obtain a clearer understanding of the dataset's overall structure, descriptive statistics were first generated. This preliminary step was essential for assessing the distribution of key variables and detecting anomalous observations that could undermine the reliability of the econometric analysis. In large-scale housing datasets, extreme values may arise from recording errors, atypical transactions, or market conditions not representative of the broader sample. Without appropriate treatment, such observations can disproportionately influence regression estimates and distort the interpretation of results.

Based on the descriptive analysis, a filtering process was implemented to exclude data points falling outside reasonable and economically plausible ranges. Specifically, transactions involving house prices below £20,000, prices per square metre below £30, and properties with floor areas smaller than 25 square metres were removed from the sample. These thresholds were established to eliminate observations that were unlikely to reflect standard residential market behaviour, such as unusually small units sold at exceptionally low prices or records affected by incomplete or erroneous reporting.

Table 1: Filtered database descriptive statistics

	Mean	Std. Dev.	Min	Max
PRICE	£466,559.96	£738,624.68	£20,000.00	£54,959,000.00
PRICE PER AREA	£5,308.62	£4,130.03	£31.25	£191,476.22
SIZE	82.98	53.64	25.00	3,055.00

Source: Made by the author

The removal of these cases serves a dual methodological purpose. First, it improves the internal consistency of the dataset by ensuring that the remaining observations correspond more closely to realistic housing market conditions. Second, it reduces the statistical leverage of extreme values that could otherwise bias coefficient estimates within the hedonic regression framework. In this sense, outlier treatment is not merely a technical adjustment, but a critical stage in preserving the validity of the analytical model.

To further enhance the comparability of observations across time, property prices were adjusted for inflation over the period from 1995 to 2021. This correction was necessary because nominal transaction values recorded over such a long timespan are affected by changes in purchasing power, rendering direct comparisons misleading if left unadjusted. By converting prices into inflation-adjusted terms, the analysis is able to assess real market trends rather than nominal increases driven solely by macroeconomic conditions. Table 2 shows the descriptive statistics adjusted for inflation.

Table 2: Inflation adjusted descriptive statistics

	Mean	Std. Dev.	Min	Max
PRICE*	£7,932.91	£995,776.02	£26,123.80	£75,226,780.02
PRICE PER AREA*	£7,932.91	£5,389.41	£56.98	£243,569.24
SIZE	£82.98	£53.64	£25.00	£3,055.00

Source: Made by the author

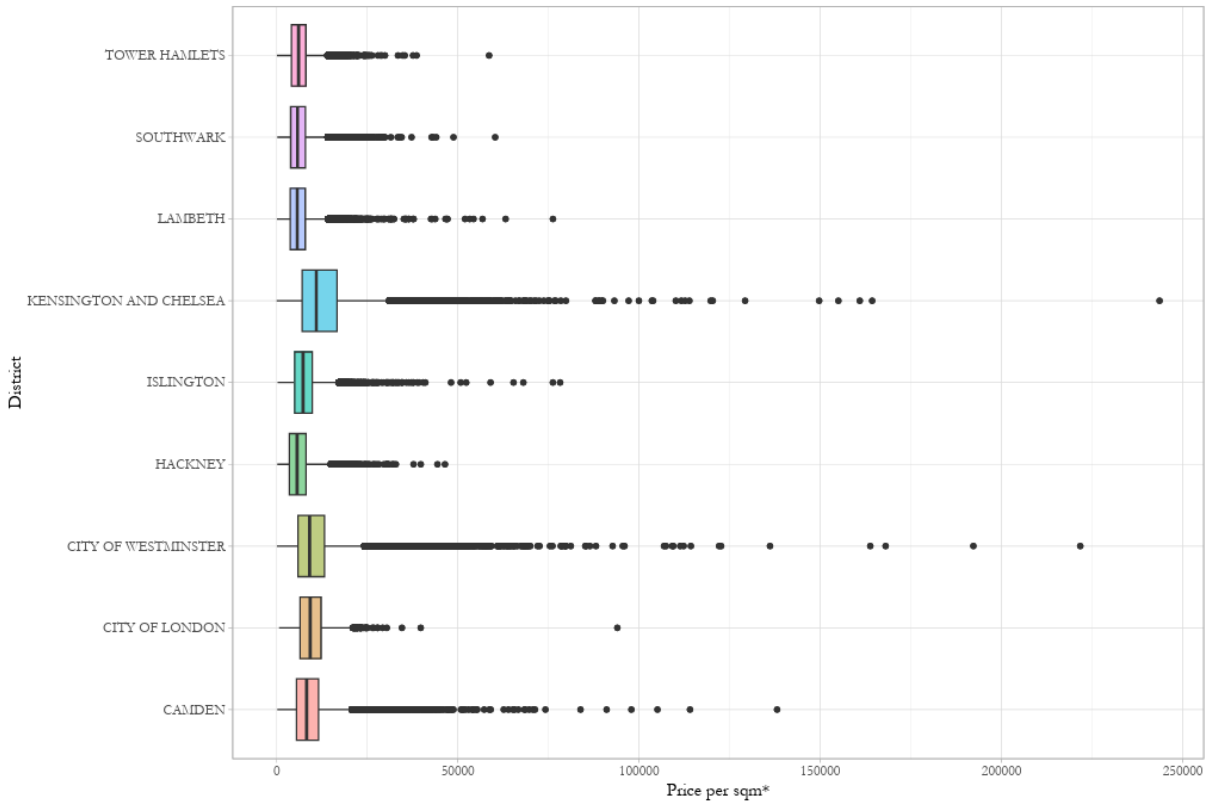
Accounting for inflation is particularly important in longitudinal housing studies, as it enables a more accurate interpretation of value appreciation and market dynamics. In the context of this research, it ensures that the estimated relationship between urban greenery and property prices reflects genuine economic patterns rather than temporal distortions. Moreover, the adjustment strengthens the relevance of the findings for present-day policy and planning considerations, as it situates historical transactions within a consistent monetary framework.

Taken together, the treatment of outliers and the inflation adjustment establish a more robust empirical basis for subsequent modelling. These procedures contribute to the reliability of the dataset and improve the capacity of the study to identify meaningful relationships between housing values and urban environmental factors.

To better understand the distribution of extreme values in housing prices, a box-and-whisker plot was developed for each neighbourhood (Figure 5). This visualisation provided an immediate comparative overview of central tendency and variability in price per square metre across districts, enabling a clearer identification of anomalies within the dataset. By representing

medians, interquartile ranges, and outlier observations, the whisker plot served as a valuable exploratory tool for assessing the heterogeneity of London's housing market.

Figure 5: Box-and-Whiskers plot of property prices per square meter divided by district



Source: Made by the author

As shown, the borough of Kensington and Chelsea exhibits a markedly higher median housing price per square metre than other districts. This is evidenced by the elevated position of the median line within a comparatively high-valued box, as well as by the presence of numerous extreme outliers extending well beyond the upper whisker. Such a pattern indicates not only a generally more expensive housing market, but also a substantial proportion of properties priced far above the standard range observed elsewhere in the city.

The concentration of outliers in this borough is not unexpected. Kensington and Chelsea is internationally recognised as one of London's most affluent residential areas, characterised by luxury housing stock, prestigious addresses, and a high concentration of premium real estate. In this context, extreme values may reflect genuinely exceptional properties distinguished by architectural significance, exclusivity, or unique locational advantages. However, it is also necessary to consider that some outliers may stem from irregular transactions, mixed-use developments, or data inconsistencies that do not accurately represent conventional residential market conditions.

From a methodological perspective, the presence of such extreme observations poses a challenge for hedonic regression analysis. Outliers exert disproportionate influence on model coefficients, potentially distorting the estimated relationship between explanatory variables and housing prices. This is particularly relevant when the research objective is to understand typical market

behaviour rather than the dynamics of ultra-luxury segments. In response, and to preserve the analytical robustness of the study, an upper cap of £35,000 per square metre was imposed on the dataset (Table 3). This threshold was selected as a pragmatic balance between retaining high-value observations that reflect legitimate market variation and excluding values that are so extreme as to risk biasing the results. By limiting the influence of exceptional cases, the analysis is better able to capture the pricing dynamics of the broader residential market across neighbourhoods.

Table 3: Inflation adjusted descriptive statistics

	Mean	Std. Dev.	Min	Max
PRICE*	£668.246,55	£831.924,65	£26.123,80	£46.949.210,00
PRICE PER AREA*	£7.803,70	£4.828,85	£56,98	£34.993,01
SIZE	82,70	52,85	25,00	3.055,00

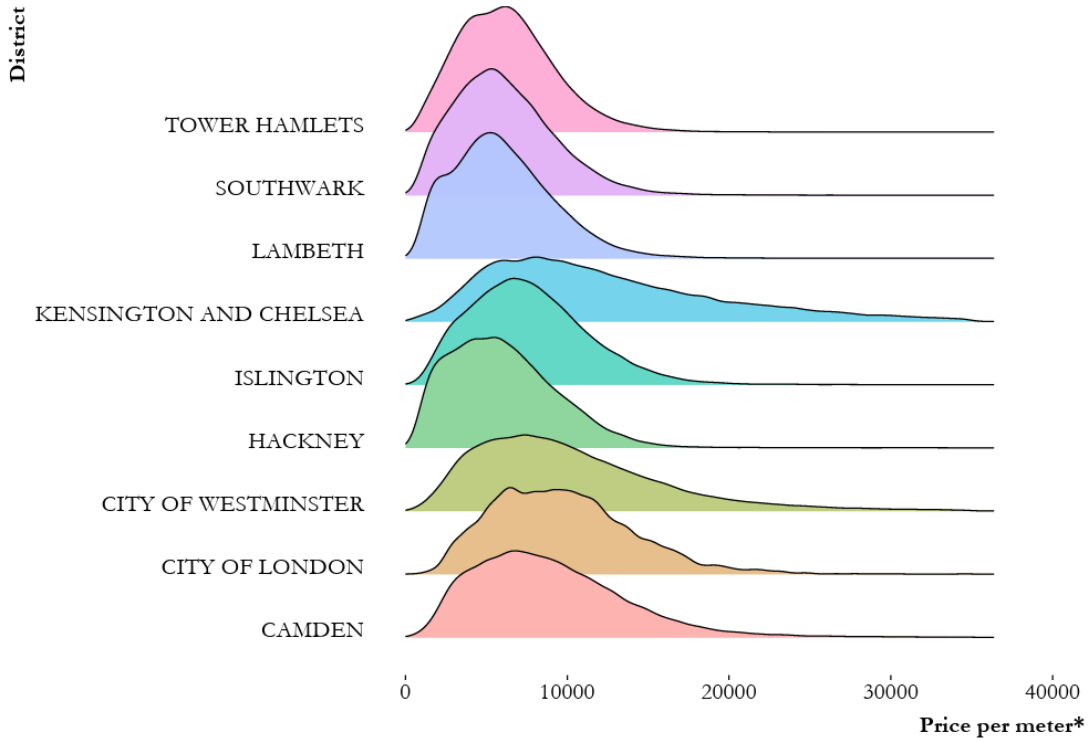
Source: Made by the author

The application of this cap does not deny the existence or relevance of ultra-premium properties; rather, it acknowledges that these transactions belong to a distinct market niche whose pricing mechanisms may differ substantially from those of the general housing stock. Excluding excessively high values ensures that the findings remain representative of typical urban residential conditions, thereby improving the interpretability and comparability of results across districts.

Ultimately, this decision strengthens the validity of the empirical framework by focusing attention on regular housing market patterns while still recognising the structural inequalities and prestige-driven pricing that characterise certain parts of London. In doing so, the research maintains its emphasis on understanding how urban amenities—particularly green spaces—affect mainstream residential values, rather than allowing rare luxury transactions to dominate the overall picture.

To complement the box-and-whisker analysis and provide a more intuitive representation of the dataset's distributional structure, a Joy-plot was developed (Figure 6). This visualisation was particularly useful for depicting the "footprint" of housing prices across districts, as it captures not only central tendency but also the overall shape and spread of the data.

Figure 6: House price distribution by district joy-plot



Source: Made by the author

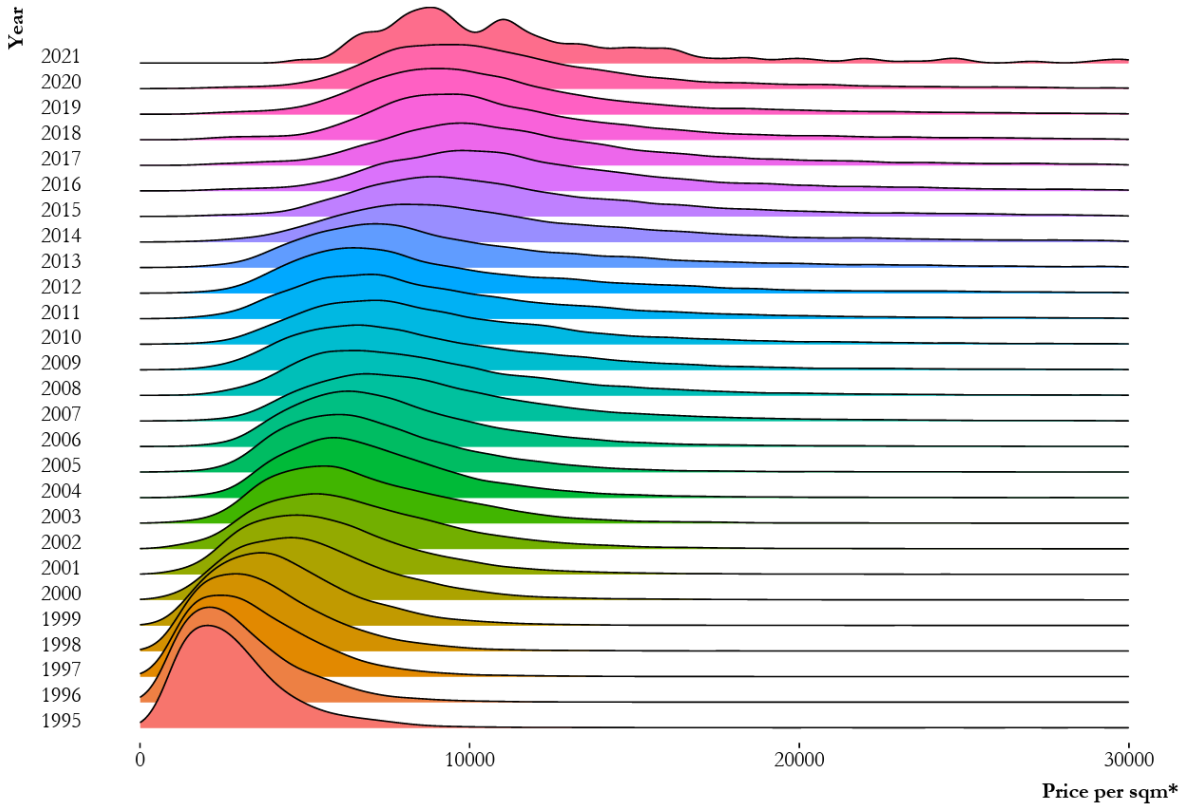
Within this framework, the peaks of the ridges correspond broadly to the median concentration of prices, while their width reflects dispersion, serving as a visual proxy for variability or standard deviation. The resulting plot confirms that Kensington and Chelsea exhibit the highest median prices per square metre, reinforcing the findings observed in the whisker analysis. By contrast, districts such as Hackney display lower price levels with more similarly distributed values, suggesting a less stratified housing market.

The shape of the ridges further reveals important structural distinctions among boroughs. Areas such as Kensington and Chelsea, the City of Westminster, and the City of London present wider and more elongated distributions, indicating substantial variation in housing prices. This suggests the coexistence of highly differentiated market segments, ranging from luxury residences to comparatively more accessible properties. Such diversity is characteristic of districts where international investment, prestige locations, and mixed residential typologies intersect.

Conversely, boroughs such as Hackney, Islington, and Camden display narrower and more concentrated distributions. This pattern implies greater consistency in housing prices and, by extension, a more homogeneous market structure. While these areas may still command significant values, the reduced dispersion suggests fewer extreme price categories and a more stable relationship between property characteristics and market valuation.

For comparative purposes, and to assess how broader market trends may influence willingness to pay for specific housing characteristics, a second Joyplot (Figure 7) was constructed to analyse the temporal dimension of price evolution. Unlike the district-based visualisation, this representation focuses on the distribution of inflation-adjusted house prices per square metre across time, thereby introducing a longitudinal perspective into the analysis.

Figure 7: House price distribution by year joyplot



Source: Made by the author

The temporal Joyplot demonstrates a pronounced long-term increase in housing prices, even when adjusted for inflation. This indicates that the upward trend cannot be attributed solely to nominal monetary expansion, but reflects a genuine rise in real estate values over the observed period. The graph also captures the temporary contraction associated with the 2008 global financial crisis, during which prices experienced a noticeable decline before resuming their upward trajectory.

Most strikingly, the final years of the dataset reveal a particularly sharp surge in prices, suggesting intensified demand and increasing scarcity within the urban housing market. This acceleration in the last five years may be interpreted as evidence of heightened market pressure, potentially linked to broader economic factors such as low interest rates, constrained housing supply, and sustained demand for centrally located urban properties.

From an analytical standpoint, these temporal shifts are highly relevant. Rising baseline prices may alter buyer sensitivity to specific property characteristics, including environmental amenities such as proximity to green spaces. In more expensive markets, willingness to pay for desirable attributes may become amplified, as households compete for increasingly scarce high-quality living environments. Conversely, rapid appreciation can also intensify affordability pressures, potentially reshaping the socio-economic composition of neighbourhoods.

The temporal Joyplot therefore provides more than a descriptive account of price growth; it contextualises the evolving economic environment within which housing preferences are

expressed. By integrating this longitudinal perspective, the research gains a deeper understanding of how market conditions may mediate the valuation of urban amenities over time.

2.6. The Hedonic Pricing Model & Hypothesis

The dependent variable selected for this study is **Price per Area**, adjusted for inflation over the observation period. This measure was chosen because it provides a standardised and economically meaningful indicator of housing value, allowing comparisons across properties of different sizes and transaction dates. Unlike total sale price, which can be heavily influenced by property scale alone, price per square metre offers a more refined metric for assessing how individual housing characteristics are capitalised into market value. By adopting this variable, the analysis is able to estimate the revealed willingness to pay associated with specific features, translating each attribute into its implicit monetary contribution.

This approach is particularly suited to hedonic regression analysis, where the objective is to decompose housing prices into the values assigned to distinct property characteristics. In practical terms, it enables the study to quantify how much more buyers are willing to pay for additional floor space, improved energy performance, enhanced accessibility, or proximity to environmental amenities such as green spaces. The dependent variable therefore acts not only as a measure of market value, but also as a lens through which consumer preferences become observable.

As discussed previously, the independent variables are organised into four broad categories of explanatory attributes, each reflecting a different dimension of residential value formation.

Structural Features include the most conventional determinants of price. These comprise **Area**, **Property Type**, **Construction Age Band**, and **Energy Rating**. Property Type is represented through dummy variables corresponding to "Flat", "House", "Maisonette", and "Bungalow", while Construction Age Band is similarly operationalised through categorical dummies ranging from "before_1900" to "2007_onwards". Energy Rating is included as a categorical measure spanning classes A to G. Together, these variables capture the physical and technical profile of each dwelling, reflecting both tangible utility and broader market perceptions of quality.

Internal Features are represented through **Glazed Areas**, operationalised via dummy variables for the categories "Much Less Than Typical", "Less Than Typical", "Normal", "More Than Typical", and "Much More Than Typical". Although limited in scope, this variable provides an important proxy for natural light availability and interior openness, both of which influence aesthetic appeal and perceived comfort.

External Features encompass locational accessibility, specifically **Proximity to Metro Stations** and **Proximity to Educational Centres**. These variables account for the extent to which integration into urban infrastructure contributes to residential desirability.

Finally, **Natural Environmental Features** are captured through **Proximity to Green Areas**, which constitutes the central explanatory variable of the study.

Collectively, these variables represent the "ingredients" that explain variations in Price per Area and allow for a more nuanced understanding of housing valuation beyond raw dimensions or aggregate transaction prices.

The hedonic pricing model is specified in (4)

$$PriceAdjPerArea = a_1(TotalFloorArea) + a_2(PropertyType) + a_3(EnergyRating) + a_4(DateOfConstruction) + a_5(GlazedArea) + a_6(District) + a_7(DistGreenery) + a_8(DistMetro) + a_9(DistEducational) + Residual \quad (4)$$

This formulation reflects the additive logic of hedonic modelling, in which each coefficient estimates the marginal contribution of a given characteristic to the overall housing price.

With regard to expected relationships, the hypothesis for **Structural Features** is that larger floor area will correlate positively with value, as increased space remains one of the strongest drivers of housing demand. Property Type is expected to produce differentiated effects, with bungalows and maisonettes potentially commanding higher premiums than houses and flats due to their relative scarcity, spatial advantages, or lifestyle appeal. Similarly, stronger Energy Ratings are anticipated to exert a positive influence, reflecting the increasing market relevance of sustainability and operational efficiency. Construction Age is also expected to matter significantly, though in complex ways: newer properties may benefit from efficiency and lower maintenance costs, while older properties may command premiums due to heritage value and architectural character.

For **Internal Features**, the expectation is likewise positive. Greater glazed areas are assumed to enhance both aesthetic appeal and perceived quality of life, particularly in dense urban settings where natural light is a valued resource.

In relation to **External Features**, proximity to metro stations and educational centres is expected to correlate positively with price, as both contribute to accessibility and long-term desirability. However, their relative importance may vary across socio-economic contexts. In lower-income areas, efficient public transport may hold greater value due to higher dependency on collective mobility. Conversely, in wealthier neighbourhoods, this effect may be less pronounced if residents rely more heavily on private transportation.

The principal hypothesis of the study concerns **Natural Environmental Features**. Specifically, it is expected that there is a positive relationship between Price per Area and proximity to green spaces. Properties located closer to parks, gardens, and other urban green areas are hypothesised to command higher prices per square metre, reflecting the premium associated with environmental quality, recreation, and improved well-being.

This assumption is grounded in the logic of revealed preferences. Rather than relying on stated attitudes, the analysis draws on actual market transactions, interpreting observed purchasing

decisions as indicators of value. As noted by Elliott et al. (2019), however, the market value of urban greenery remains contingent upon human perception and awareness. Green spaces generate benefits only insofar as these are recognised and internalised by buyers, making their valuation both economic and socially constructed.

An important nuance is that the effect of proximity to green areas may not be linear. Its significance could diminish beyond a certain threshold, or vary according to neighbourhood income levels. In lower-priced areas, affordability may outweigh environmental considerations, reducing willingness to pay for nearby greenery. In contrast, in affluent districts where basic housing needs are already secured, environmental amenities may become more salient and command stronger premiums.

A similar inverse logic may apply to transport accessibility: wealthier households may display lower dependence on metro proximity, while lower-income groups may place greater emphasis on public transport connections. These interactions suggest that the valuation of amenities is mediated by socio-economic context, rather than universally fixed.

Finally, the impact of green areas is likely to depend not only on distance, but also on the size and type of the green space itself, as highlighted by Kong et al. (2007) and Kim et al. (2019). A large park equipped with walking paths, playgrounds, sports facilities, or event spaces may generate substantially greater premiums than a small landscaped corner or modest green patch. Such distinctions underscore the importance of recognising that not all greenery is valued equally. The quality, functionality, and symbolic meaning of green spaces can significantly shape their contribution to residential desirability and, consequently, to housing prices.

3. *EMPIRICAL RESULTS*

In order to assess the robustness of the estimated effects, a sequence of nested hedonic regression models was estimated. Rather than relying solely on the fully specified model, explanatory variables were introduced progressively according to the conceptual framework developed in the previous chapter (see Table 4). This approach allows the stability of the estimated coefficients to be evaluated and provides evidence regarding the extent to which the observed relationships depend on the inclusion of specific control variables.

Table 4: Model Content

Model	**Features included**
Model 1	Structural features only
Model 2	Structural + Internal Feature GlazedArea
Model 3	Structural + Internal + External Features
Model 4	Structural + Internal + External Features + Greenery Proximity

Source: Made by the author

Comparing the estimated coefficients across specifications allows an assessment of robustness. If the coefficients remain relatively stable as additional controls are introduced, this suggests that the relationships identified are not driven by omitted variables and can be considered robust. Particular attention is given to the coefficient associated with Dist_Greenery, as the central objective of this study is to evaluate the extent to which proximity to green spaces is capitalised into housing prices.

Beginning the empirical analysis, the results reveal a substantial improvement in explanatory power as additional variables are incorporated into the analysis whose results are shown in Table 5.

Table 5: Model Summary

Model	Residual Std. Error	Residual DF	R^2	Adjusted R^2	F-statistic	Model DF	Model p-value
Model 1	4,685	472,402	0.05868	0.05864	1,472	20	<2.2e-16
Model 2	4,683	472,398	0.05950	0.05946	1,245	24	<2.2e-16
Model 3	4,224	472,388	0.23500	0.23500	4,269	34	<2.2e-16
Model 4	4,224	472,387	0.23510	0.23500	4,147	35	<2.2e-16

Source: Made by the author

Model 1, which includes only structural characteristics, explains approximately 5.9% of the variation in inflation-adjusted price per square metre (R^2 = 0.0587). Although this relatively modest explanatory power is expected, it indicates that structural characteristics alone are insufficient to capture the complexity of housing prices in London.

The inclusion of GlazedArea in Model 2 produces only a marginal improvement in model fit, increasing the R^2 from 0.0587 to 0.0595. This suggests that internal glazing characteristics contribute positively to explaining property prices, although their independent contribution remains relatively limited compared with broader structural and locational factors.

A much larger improvement is observed in Model 3 after introducing external variables and district effects. The R^2 increases to 0.2350, meaning that approximately 23.5% of the variation in housing prices is explained by the model. Simultaneously, the Residual Standard Error declines substantially from approximately 4,683 to 4,224, indicating a considerably better fit between predicted and observed prices. This finding confirms that location plays a dominant role in determining residential property values within London.

The final specification, Model 4, incorporates the proximity to green spaces variable. The explanatory power increases only marginally from 0.2350 to 0.2351, while the Residual Standard Error remains virtually unchanged. Although the increase in R^2 is relatively small, this outcome should not be interpreted as evidence that green spaces are unimportant. Rather, it indicates that green space proximity explains a relatively small but statistically significant component of housing price variation after controlling for numerous structural and locational characteristics.

Importantly, all four models are highly statistically significant, with F-statistics ranging from 1,245 to 4,269 and p-values below 2.2×10^{-16}, confirming that the explanatory variables jointly explain a significant proportion of housing price variation.

The residual summary further demonstrates the progressive improvement of model performance as shown in Table 6.

Table 6: Residuals Summary from the regression models

Model	Min	1Q	Median	3Q	Max
Model 1	-36,494.0	-3,108.0	-830.0	2,029.0	28,992.0
Model 2	-35,780.0	-3,107.0	-829.0	2,029.0	29,015.0
Model 3	-14,883.0	-2,731.2	-450.2	2,085.5	28,578.6
Model 4	-14,877.1	-2,732.4	-450.3	2,086.4	28,610.8

Source: Made by the author

Models 1 and 2 exhibit considerably larger residual dispersion, with minimum residuals below -35,000 and median residuals close to -830. After introducing accessibility variables in Model 3, both the minimum residual and the median residual improve substantially, while the interquartile range becomes noticeably narrower.

Model 4 produces almost identical residual statistics to Model 3, suggesting that the addition of green space proximity improves the explanatory capacity of the model without substantially altering its overall predictive accuracy. The stability of the residual distribution indicates that the final specification remains well behaved after including the environmental variable.

Furthermore, one of the principal objectives of estimating nested models is to evaluate the stability of the estimated coefficients across alternative specifications, shown in Table 7, Table 8, Table 9, and Table 10.

Table 7: Coefficient estimates by model

Variable	Model 1 Estimate	Model 2 Estimate	Model 3 Estimate	Model 4 Estimate
Intercept	7,743.905	7,994.655	9,883.000	9,926.000
TOTAL_FLOOR_AREA	9.101	8.867	0.820	0.823
Property_TypeFlat	281.156	296.046	-1,122.000	-1,123.000
Property_TypeHouse	-545.139	-531.266	-641.100	-643.900
Property_TypeMaisonette	-1,204.812	-1,196.501	-1,727.000	-1,728.000
Energy_RatingB	-609.822	-536.857	290.100	289.400
Energy_RatingC	-688.728	-622.273	227.200	226.700
Energy_RatingD	-960.589	-898.702	169.100	168.600
Energy_RatingE	-1031.831	-971.011	155.900	155.200

Energy_RatingF	-599.490	-543.124	399.500	398.700
Energy_RatingG	-670.656	-616.653	241.800	240.700
Date_of_Construction1930_1949	-222.452	-219.971	-479.600	-479.400
Date_of_Construction1950_1966	-1,172.750	-1,183.117	-1,080.000	-1,081.000
Date_of_Construction1967_1975	-1,157.484	-1,182.916	-1,296.000	-1,298.000
Date_of_Construction1976_1982	-904.369	-913.008	-972.500	-973.500
Date_of_Construction1983_1990	-816.238	-829.980	-299.100	-302.300
Date_of_Construction1991_1995	-960.023	-974.100	-319.100	-321.400
Date_of_Construction1996_2002	-315.990	-351.635	444.800	443.600
Date_of_Construction2003_2006	804.673	746.269	1,668.000	1,668.000
Date_of_Construction2007_onwards	1,637.504	1,570.808	2,471.000	2,469.000
Date_of_Constructionbefore_1900	1,453.003	1,450.345	528.800	529.600
GlazedAreaMore Than Typical	—	372.008	579.300	578.900
GlazedAreaMuch Less Than Typical	—	-190.779	545.300	549.600
GlazedAreaMuch More Than Typical	—	144.077	399.000	399.400
GlazedAreaNormal	—	-328.577	-35.300	-35.100
DistrictCity of London	—	—	879.300	865.600
DistrictCity of Westminster	—	—	1,123.000	1,117.000
DistrictHackney	—	—	-2382.000	-2383.000
DistrictIslington	—	—	-1,349.000	-1,362.000
DistrictKensington and Chelsea	—	—	3,164.000	3,161.000
DistrictLambeth	—	—	-2,455.000	-2,443.000
DistrictSouthwark	—	—	-2,220.000	-2,222.000
DistrictTower Hamlets	—	—	-2,681.000	-2,696.000
DIST_METRO	—	—	-0.435	-0.437
DIST_EDUCATIONAL	—	—	0.153	0.156
Dist_Greenery	—	—	—	-0.204

Source: Made by the author

Table 8: Main p-values by model

Variable	Model 1 p-value	Model 2 p-value	Model 3 p-value	Model 4 p-value
Intercept	9.20e-11	2.49e-11	<2e-16	<2e-16
TOTAL_FLOOR_AREA	<2e-16	<2e-16	3.29e-09	2.96e-09
Property_TypeFlat	0.2320	0.207973	1.24e-07	1.20e-07
Property_TypeHouse	0.0207	0.024053	0.00254	0.00243
Property_TypeMaisonette	3.35e-07	3.98e-07	5.10e-16	4.78e-16
Energy_RatingB	0.6028	0.646706	0.78358	0.78413
Energy_RatingC	0.5566	0.595173	0.82965	0.83008
Energy_RatingD	0.4123	0.442861	0.87280	0.87316
Energy_RatingE	0.3785	0.407077	0.88265	0.88317
Energy_RatingF	0.6091	0.643005	0.70539	0.70595
Energy_RatingG	0.5681	0.599527	0.81945	0.82025
GlazedAreaMore Than Typical	—	0.000104	2.06e-11	2.11e-11
GlazedAreaMuch Less Than Typical	—	0.602174	0.09856	0.09594
GlazedAreaMuch More Than Typical	—	0.202506	9.19e-05	9.04e-05
GlazedAreaNormal	—	0.000223	0.66018	0.66198
DIST_METRO	—	—	<2e-16	<2e-16
DIST_EDUCATIONAL	—	—	1.42e-11	5.08e-12
Dist_Greenery	—	—	—	6.09e-05

Source: Made by the author

Table 9: ANOVA — F values

Variable	Model 1 F value	Model 2 F value	Model 3 F value	Model 4 F value
TOTAL_FLOOR_AREA	4,050.274	4,053.780	4,983.871	4,984.030
Property_Type	2,079.737	2,081.540	2,559.121	2,559.203
Energy_Rating	75.864	75.930	93.351	93.354
Date_of_Construction	1,870.457	1,872.080	2,301.602	2,301.675
GlazedArea	—	103.360	127.077	127.081
District	—	—	13,126.323	13,126.742
DIST_METRO	—	—	3,339.173	3,339.279
DIST_EDUCATIONAL	—	—	45.649	45.650
Dist_Greenery	—	—	—	16.075

Source: Made by the author

Table 10: ANOVA — P values

Variable	Model 1 p-value	Model 2 p-value	Model 3 p-value	Model 4 p-value
TOTAL_FLOOR_AREA	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
Property_Type	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
Energy_Rating	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
Date_of_Construction	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
GlazedArea	—	<2.2e-16	<2.2e-16	<2.2e-16
District	—	—	<2.2e-16	<2.2e-16
DIST_METRO	—	—	<2.2e-16	<2.2e-16
DIST_EDUCATIONAL	—	—	1.417e-11	1.415e-11
Dist_Greenery	—	—	—	6.090e-05

Source: Made by the author

The coefficient associated with Floor Area remains positive throughout all four models, confirming that larger dwellings command higher prices per square metre. While the estimated coefficient decreases considerably after locational controls are introduced—from approximately 9.1 in Model 1 to approximately 0.82 in Models 3 and 4—this change reflects the fact that part of the apparent size premium is actually explained by differences in neighbourhood characteristics. Once these confounding influences are controlled for, the estimated marginal effect of floor area stabilises.

Similarly, the coefficients associated with Property Type remain relatively stable once accessibility variables are introduced. Flats, houses and maisonettes all exhibit lower prices relative to the omitted reference category, with maisonettes displaying the largest negative premium. The consistency of these estimates across Models 3 and 4 suggests that housing typology exerts an independent influence on market value.

Energy Rating remains statistically insignificant throughout all specifications. Although the coefficients change sign after locational controls are introduced, none of the estimated effects achieve conventional significance levels. This finding suggests that energy efficiency does not independently influence housing prices within the present specification after accounting for other structural and locational characteristics.

Construction period consistently represents one of the strongest determinants of property values. Recently constructed properties (particularly those built after 2003) exhibit substantial positive premiums, while many mid-twentieth-century properties display negative coefficients relative to the reference category. These estimates remain remarkably stable across Models 3 and 4, providing additional evidence of coefficient robustness.

The inclusion of GlazedArea in Models 2–4 also reveals relatively stable effects. Properties with more than typical or much more than typical glazing consistently command positive and statistically significant price premiums, suggesting that natural light and openness are valued by homebuyers.

Perhaps the most notable change occurs after district effects and accessibility variables are introduced in Model 3. District coefficients become the largest predictors within the model, confirming that neighbourhood location represents one of the primary determinants of housing prices. Kensington and Chelsea exhibit the largest positive premium, while Hackney, Lambeth, Southwark and Tower Hamlets display negative coefficients relative to the reference district. Likewise, proximity to metro stations exerts a strong negative coefficient, indicating that increasing distance from public transport significantly reduces property values.

Finally, Model 4 introduces the central variable of the study, Dist_Greenery. The coefficient is negative (-0.204) and highly statistically significant ($p = 6.09 \times 10^{-5}$), indicating that, after controlling for structural characteristics, internal features, district effects, and accessibility variables, each additional metre of distance from green spaces is associated with a reduction of approximately **£0.20/m²** in housing value. Equivalently, an additional 100 m distance from green infrastructure corresponds to an estimated reduction of approximately **£20/m²**. Although the magnitude of the coefficient is relatively modest compared with major structural determinants such as district location, the statistical significance confirms that environmental amenities contribute independently to housing valuation.

Looking into the sequential ANOVA results provide additional evidence regarding the contribution of each group of explanatory variables. Across all four models, Total Floor Area, Property Type, Energy Rating, and Date of Construction remain highly significant, with p-values below 2.2×10^{-16}.

The inclusion of GlazedArea in Model 2 yields an F-value exceeding 103, increasing to approximately 127 in Models 3 and 4, indicating that internal housing characteristics contribute significantly to explaining price variation.

The introduction of district effects in Model 3 produces the highest F-value (13,126), confirming that geographical location represents the single most influential determinant of housing prices within the study area. Distance to metro stations also exhibits a very large F-value (approximately 3,339), reinforcing the importance of transport accessibility in London's residential market.

Finally, Dist_Greenery records an F-value of 16.08 with a p-value of 6.09×10^{-5}. Although considerably smaller than the values associated with district or transport accessibility, this result nevertheless confirms that proximity to green spaces makes a statistically significant contribution to explaining housing prices after accounting for all other explanatory variables.

Variance Inflation Factor (VIF) diagnostics indicate that multicollinearity is not a concern in any of the estimated models. All adjusted GVIF values remain well below conventional thresholds, with the highest value equal to approximately 1.27 for the metro accessibility variable. Even after introducing district effects and green space proximity, no substantial increase in multicollinearity is observed.

These findings shown in Table 11 suggest that the explanatory variables provide largely independent information regarding housing prices, thereby supporting the reliability and stability of the estimated regression coefficients.

Table 11: VIF — Adjusted GVIF values

Variable	Model 1	Model 2	Model 3	Model 4
TOTAL_FLOOR_AREA	1.168350	1.172433	1.192469	1.192480
Property_Type	1.064742	1.064794	1.077682	1.077770
Energy_Rating	1.041980	1.042118	1.042859	1.042861
Date_of_Construction	1.024332	1.025069	1.036910	1.036997
GlazedArea	—	1.003158	1.003837	1.003840
District	—	—	1.048381	1.054540
DIST_METRO	—	—	1.269057	1.272899
DIST_EDUCATIONAL	—	—	1.006228	1.006950
Dist_Greenery	—	—	—	1.055349

Source: Made by the author

Overall, the progressive estimation strategy demonstrates a high degree of robustness. As additional variables are introduced, the principal coefficients remain relatively stable and retain their expected signs. The inclusion of accessibility controls substantially improves explanatory

power, while the addition of the green space variable introduces a statistically significant environmental effect without materially altering the estimates of the remaining coefficients.

Consequently, the nested regression analysis provides strong evidence that the positive valuation of proximity to green spaces is not driven by omitted structural or locational characteristics. Rather, green infrastructure appears to generate an independent, although comparatively modest, contribution to residential property values within the London housing market.

In addition, a further dimension of the analysis explores the role of sociodemographic and neighbourhood-level heterogeneity, following an approach similar to that of Su et al. (2021). To this end, a disaggregated analysis at the district level (Table 12) was conducted in order to assess whether different urban communities exhibit varying sensitivities to green space accessibility. This specification is designed to capture potential heterogeneity in preferences for green amenities, which may be shaped by income levels, lifestyle characteristics, housing demand structures, and broader socio-cultural factors influencing residential choice and valuation processes.

Table 12: District Analysis

District	Significant Predictors	R-squared	F-statistic	p-value
Camden	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, DIST_METRO, Dist_Greenery	0.04837	87.65	< 2.2e-16
City of London	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, Dist_Greenery	0.04476	7.294	< 2.2e-16
City of Westminster	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, Dist_Greenery	0.05496	149.3	< 2.2e-16
Hackney	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, DIST_METRO, Dist_Greenery	0.08881	145.3	< 2.2e-16
Islington	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, Dist_Greenery	0.0686	121.2	< 2.2e-16
Kensington and Chelsea	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, GlazedArea, Dist_Greenery	0.05274	105.1	< 2.2e-16
Lambeth	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, DIST_METRO, Dist_Greenery	0.1302	434.9	< 2.2e-16
Southwark	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, DIST_METRO, Dist_Greenery	0.08396	249.1	< 2.2e-16
Tower Hamlets	TOTAL_FLOOR_AREA, Property_Type, Date_of_Construction, DIST_METRO, Dist_Greenery	0.1918	609	< 2.2e-16

Source: Made by the author

At the district level, the model results indicate variation in explanatory power across London's boroughs, as reflected in differing R-squared values. Camden, the City of London, and the City of Westminster display relatively low R-squared values of 0.04837, 0.04476, and 0.05496 respectively, suggesting that while the models are statistically significant, a large proportion of price variation remains unexplained in these areas. This may reflect the high degree of market complexity in central London districts, where property values are influenced by highly

heterogeneous factors such as exclusivity premiums, international demand, and micro-location effects that are not fully captured by the current specification.

By contrast, Tower Hamlets exhibits a notably higher R-squared value of 0.1918, indicating that approximately 19.18% of the variation in housing prices is explained by the model in this district. This suggests a comparatively stronger alignment between the selected predictors and observed price dynamics in this area, potentially reflecting more structured or homogeneous pricing determinants relative to more central or highly segmented markets.

Across all districts, the F-statistics are consistently high—for example, 87.65 in Camden, 149.3 in the City of Westminster, and 609 in Tower Hamlets—with p-values below 2.2e-16 in all cases. This confirms that, at the district level, the regression models are jointly statistically significant and that at least one predictor meaningfully contributes to explaining variation in housing prices within each borough.

Regarding multicollinearity, the VIF values remain consistently low across all district-level models, with all reported values below conventional thresholds (e.g., below 2 in Camden). This indicates that multicollinearity does not pose a significant threat to the reliability of coefficient estimates, and that the explanatory variables retain sufficient independence to allow for stable and interpretable regression results.

In addition, it is noteworthy that districts such as Lambeth and Tower Hamlets exhibit comparatively higher R-squared values relative to other boroughs, suggesting that the selected explanatory variables have stronger predictive power in these areas. This is further supported by the magnitude of the corresponding F-statistics, which confirm the overall statistical robustness of the models. Such variation across districts underscores the heterogeneous nature of London's housing market, where structural, locational, and environmental factors interact differently depending on the socio-spatial characteristics of each borough.

Now, turning the attention to the topic variable, a key observation emerging from the district-level regressions is that Dist_Greenery remains statistically significant across most specifications (Table 13). This reinforces the interpretation that proximity to substantial green infrastructure plays a meaningful role in shaping housing prices, especially within highly urbanised environments such as London. The persistence of this significance across multiple districts suggests that the value attributed to green amenities is not confined to a single neighbourhood type but rather represents a broader structural feature of urban housing valuation.

Table 13: District Analysis for Distance to Greenery

District	Estimate (£/m² per metre distance)	Std. Error	t value	p-value	Significance	Interpretation
Camden	-0.3446	0.1504	-2.292	0.021925	*	Distance from greenery decreases property value
City of London	-0.4897	0.7181	-0.682	0.495277	Not significant	No statistical evidence of a greenery effect
City of Westminster	0.2964	0.2187	1.355	0.175369	Not significant	No statistical evidence of a greenery effect
Hackney	-0.5943	0.1315	-4.52	6.19E-06	***	Distance from greenery significantly decreases property value
Islington	-1.582	0.1741	-9.089	<2e-16	***	Strong negative effect of distance from greenery
Kensington and Chelsea	0.4534	0.1615	2.806	0.005011	**	Unexpected positive relationship with distance
Lambeth	-0.5981	0.07475	-8.002	1.24E-15	***	Strong negative effect of distance from greenery
Southwark	0.7407	0.08259	8.969	<2e-16	***	Unexpected positive relationship with distance
Tower Hamlets	1.368	0.1183	11.565	<2e-16	***	Strong positive relationship with distance

Source: Made by the author

The strongest evidence of a positive valuation of green proximity is observed in districts such as **Islington, Lambeth, Hackney, and Camden**, where increasing distance from green areas is associated with lower property prices. These results suggest that in these neighbourhoods, parks and green spaces function as valuable residential amenities, potentially improving perceived neighbourhood quality, environmental comfort, and recreational accessibility. In these areas, proximity to greenery appears to be integrated into buyers' preferences and is reflected in market prices after controlling for structural, locational, and property characteristics.

The particularly strong effect observed in **Islington and Lambeth** suggests that green infrastructure may act as a scarce amenity in dense urban environments, where access to open space is limited and therefore more highly valued. In such contexts, properties located close to parks may benefit from increased attractiveness due to improved living conditions, visual amenity, recreational opportunities, and potentially lower exposure to urban pressures such as traffic and congestion.

However, the absence of a significant relationship in **City of London and City of Westminster** indicates that in highly central areas, the influence of greenery may be overshadowed by other dominant determinants of property value, such as accessibility, employment concentration,

prestige location, historical character, and proximity to major urban functions. In these markets, green spaces may represent an additional amenity but not a primary driver of residential price differentiation.

The positive coefficients observed in **Tower Hamlets, Southwark, and Kensington and Chelsea** should not be interpreted as evidence that greater distance from parks increases property desirability. Instead, they likely reflect more complex spatial patterns, where premium residential locations may be concentrated away from the selected green spaces, or where the characteristics of the considered parks do not fully represent the green amenities valued by residents. This highlights the importance of defining green infrastructure appropriately, as the measured effect depends on which spaces are included in the analysis and how accessibility to greenery is represented.

In conclusion, while R-squared values remain relatively modest in several districts—indicating that additional unobserved factors likely influence housing prices—the consistent statistical significance of key variables, particularly Dist_Greenery, suggests that both environmental and structural characteristics play an important role in shaping property values. Overall, the results indicate that green infrastructure contributes to housing valuation, but its influence is highly dependent on neighbourhood context. The strongest impacts are found where green spaces represent a relatively scarce and valued residential amenity, while in central or highly heterogeneous districts, their effect becomes weaker or may be captured by broader spatial factors. The results suggest that the perceived value of greenery is influenced not only by distance but also by the scale and relevance of the green space itself available to residents.

4. LIMITATIONS AND FUTURE RESEARCH

Although this research provides insights into the relationship between urban green infrastructure and residential property values in London, several limitations should be acknowledged. These limitations primarily relate to the methodological assumptions of the hedonic pricing approach, the measurement of environmental amenities, the characteristics of the available dataset, and the complexity of housing market dynamics. Recognising these constraints is essential for correctly interpreting the findings and identifying opportunities for future research.

4.1. Methodological limitations

The first limitation concerns the use of the hedonic pricing method itself. As discussed throughout this research, hedonic regression provides a valuable framework for estimating the implicit value of characteristics that are not directly traded in markets, such as access to green infrastructure. By analysing observed housing transactions, the method allows researchers to identify whether environmental amenities are reflected in residential prices. However, the approach relies on several assumptions that may not be fully satisfied within complex urban housing markets.

One important assumption is that housing characteristics can be separated into individual components whose contributions to price can be independently estimated. In reality, residential preferences are shaped by the interaction of multiple characteristics rather than isolated attributes. The value associated with a property may emerge from the combination of location, architectural quality, neighbourhood identity, accessibility, environmental conditions, and social context. Consequently, the estimated coefficient associated with proximity to green infrastructure should not be interpreted as a complete measure of the value of urban greenery, but rather as the marginal contribution of this characteristic within the specific model specification adopted.

Furthermore, hedonic regression identifies associations between observed characteristics and market prices but does not necessarily establish causal relationships. Although the model controls for several important determinants of housing values, unobserved characteristics may still influence both the location of green spaces and residential prices. For example, historically desirable neighbourhoods may simultaneously contain higher-quality housing, better maintained public spaces, and higher property values. In such cases, part of the estimated effect associated with greenery may reflect broader neighbourhood characteristics rather than the direct influence of green infrastructure itself.

Future research could address this limitation through the application of more advanced econometric approaches, such as quasi-experimental designs, longitudinal analysis, or instrumental variable strategies using stronger external instruments. These methods could

provide further evidence regarding the causal mechanisms through which green infrastructure influences housing markets.

4.2. Limitations in measuring green infrastructure

A second limitation relates to the representation of urban greenery within the model. In this research, access to green infrastructure is measured through distance from residential properties to the nearest significant green space. This indicator provides a clear and replicable measure of accessibility, allowing the relationship between location and housing values to be statistically examined. However, distance alone does not fully capture the complexity of how residents experience and evaluate green spaces.

Urban green areas differ substantially in terms of size, ecological quality, maintenance, accessibility, safety, recreational opportunities, and aesthetic characteristics. A large metropolitan park and a small neighbourhood garden may both be classified as green spaces, but they provide different types and levels of benefits. Similarly, proximity does not necessarily imply effective access, as physical barriers, transport connections, opening hours, or perceived safety may influence whether residents actually use these spaces.

The findings of this research should therefore be interpreted as reflecting the economic value associated with spatial proximity to green infrastructure rather than the complete value of all ecosystem and social services provided by urban greenery.

Future studies could improve this aspect by incorporating multidimensional indicators of green infrastructure. These could include measures of green space size, accessibility networks, vegetation quality, biodiversity indicators, recreational facilities, and residents' perceptions of environmental quality. Combining objective spatial measures with subjective assessments could provide a more comprehensive understanding of how green amenities are valued by different communities.

4.3. Dataset limitations and omitted variables

A further limitation concerns the characteristics of the available housing dataset. The dataset used in this research provides extensive information on property transactions, structural characteristics, energy performance, and spatial attributes, creating a strong empirical foundation for hedonic analysis. Nevertheless, administrative housing datasets generally contain limited information regarding qualitative aspects of residential properties.

Important characteristics such as interior condition, architectural design, renovation quality, views, noise exposure, neighbourhood reputation, and detailed building management are not fully represented. These factors may influence both property prices and residential preferences, potentially affecting the estimated relationship between green infrastructure and housing values.

Although the inclusion of multiple structural and locational variables reduces the risk of omitted variable bias, it cannot completely eliminate it. Housing markets are influenced by complex social, economic, and cultural factors that are difficult to measure quantitatively. In particular, household income, demographic characteristics, and individual preferences may affect both residential location choices and the perceived value of environmental amenities.

Future research could therefore benefit from integrating additional datasets containing socio-economic information, property quality indicators, and household-level characteristics. Combining housing transaction data with surveys or qualitative research could also provide insights into why residents value green spaces and how these preferences differ between population groups.

4.4. Spatial and contextual limitations

The geographical focus of this research represents both a strength and a limitation. London provides an exceptional case study due to its diverse housing market, extensive green infrastructure network, and availability of high-quality spatial data. The variation between neighbourhoods allows for an investigation of how environmental amenities interact with different urban contexts.

However, the findings are specific to the selected areas of London and should not automatically be generalised to other cities or urban environments. The relationship between green infrastructure and housing values depends heavily on local conditions, including urban density, housing market pressure, cultural preferences, planning history, and the availability of alternative amenities.

Even within London itself, different boroughs may experience different relationships between environmental amenities and property values. The weaker influence of greenery in some central locations may reflect the dominance of other urban advantages, such as employment accessibility, historical significance, or transport connectivity. Conversely, in areas where green space is scarce, environmental amenities may represent a stronger differentiating factor.

Future research could extend this analysis by comparing different metropolitan contexts or by examining variations between inner-city, suburban, and peripheral areas. Comparative studies across cities would provide a clearer understanding of whether the patterns observed in London represent broader trends or are the result of specific local characteristics.

4.5. Future research directions

Building upon the limitations identified above, several directions for future research emerge.

First, future studies could investigate green infrastructure through a more detailed classification of different green space types. Rather than treating parks and gardens as a single category,

research could examine whether specific forms of greenery generate different economic and social outcomes. Large metropolitan parks, community gardens, urban forests, and ecological corridors may provide distinct benefits and therefore produce different patterns of capitalisation.

Second, further research could explore the relationship between green infrastructure investment and housing affordability. While increased property values may indicate that residents value environmental amenities, rising prices may also contribute to displacement pressures and unequal access to urban improvements. Understanding this relationship is particularly important for cities experiencing simultaneous environmental and housing challenges.

Third, future studies could examine the temporal dimension of green infrastructure valuation. Housing markets evolve over time, and the introduction of new parks, regeneration projects, or environmental improvements may influence property values differently depending on when and where interventions occur. Longitudinal studies could therefore provide stronger evidence regarding how environmental investments reshape urban markets.

Finally, integrating economic, environmental, and social perspectives would provide a more complete understanding of green infrastructure value. The contribution of urban greenery cannot be reduced exclusively to its impact on property prices, as many benefits—such as improved well-being, climate adaptation, and biodiversity enhancement—are not directly captured by market transactions. Future research should therefore continue developing interdisciplinary approaches capable of recognising both the measurable and non-market values of urban nature.

CONCLUSION

This research investigated the relationship between urban green amenities and residential property values in London through the application of a hedonic pricing model. The study was motivated by the increasing recognition that green infrastructure represents an essential component of contemporary urban environments, providing environmental, social, and economic benefits. While the ecological and social functions of urban greenery are widely acknowledged, the extent to which these benefits are incorporated into housing markets remains a complex and context-dependent question.

By analysing residential transactions within selected areas of central London, this research examined whether proximity to major green spaces contributes to variations in housing prices after controlling for a range of structural, building-quality, and accessibility characteristics. The analysis incorporated variables related to property characteristics, energy performance, transport accessibility, educational facilities, and distance from green infrastructure, allowing the contribution of environmental amenities to be evaluated within the broader structure of urban value formation.

The findings indicate that urban greenery represents a measurable component of residential value; however, its influence is neither uniform nor independent from other determinants of housing prices. Consistent with previous research on hedonic valuation, the results confirm that structural characteristics, particularly floor area, property type, and construction characteristics, remain among the strongest predictors of residential value. Similarly, locational factors related to accessibility and neighbourhood conditions play a fundamental role in shaping housing market outcomes.

Within this broader framework, proximity to green spaces contributes additional explanatory power, suggesting that environmental amenities are recognised by the housing market and may be partially capitalised into property prices. However, the magnitude of this effect is relatively limited when compared with major structural and accessibility variables. This highlights that green infrastructure should not be interpreted as a standalone determinant of housing value, but rather as one element within a complex network of characteristics that collectively influence residential preferences and market outcomes.

A central contribution of this research is the demonstration that the economic value of green infrastructure is spatially dependent. The district-level analysis reveals that the relationship between greenery and housing values varies considerably across London's urban contexts. While some areas show a clearer association between proximity to green spaces and higher property values, others display weaker relationships, indicating that the role of environmental amenities changes according to local market conditions, urban density, existing accessibility patterns, and the availability of alternative amenities.

These findings reinforce the argument developed throughout the literature review that green spaces should not be considered homogeneous assets. Their contribution to residential

desirability depends not only on their physical presence but also on their accessibility, perceived quality, surrounding urban functions, and the characteristics of the communities they serve. A large metropolitan park, a neighbourhood garden, and a small landscaped area may all provide environmental benefits, but their social and economic significance differs according to their position within the urban system.

The results therefore suggest that the value of green infrastructure extends beyond a simple relationship between distance and price. Although proximity provides a useful measurable indicator, it represents only one dimension of the broader relationship between residents and urban nature. The appreciation of green spaces is also influenced by how these spaces are experienced, perceived, and integrated into everyday urban life. In this sense, the capitalisation of greenery into housing prices reflects not only physical accessibility but also the wider social and environmental qualities associated with particular neighbourhoods.

From a planning perspective, these findings highlight the importance of considering green infrastructure as a multifunctional element of urban development. While economic appreciation of green amenities can demonstrate their relevance within housing markets, the value of green spaces should not be reduced solely to their capacity to generate property premiums. Urban greenery contributes to climate adaptation, biodiversity, public health, recreational opportunities, and social well-being, benefits that are not entirely captured through market transactions.

Furthermore, the relationship between green infrastructure and housing values requires careful consideration regarding issues of accessibility and equity. If improvements in environmental quality contribute to increasing property values, they may also intensify affordability pressures in already competitive housing markets. Therefore, planning strategies should aim not only to increase the provision of green infrastructure but also to ensure that its benefits are distributed inclusively across different communities and neighbourhoods.

The methodological approach adopted in this research provides valuable insights into the economic dimension of urban greenery, while also presenting several limitations. The use of distance from green spaces as an indicator captures accessibility but does not fully account for differences in size, quality, maintenance, biodiversity, functionality, or residents' perceptions of individual green areas. Additionally, although the model incorporates a wide range of structural and locational variables, housing markets are influenced by numerous characteristics that are difficult to observe, including architectural quality, internal conditions, neighbourhood identity, and household preferences.

Future research could therefore extend this analysis by incorporating more detailed measures of green infrastructure quality and accessibility. Rather than relying exclusively on proximity, future studies could examine indicators related to park size, usability, ecological characteristics, recreational facilities, and perceived environmental quality. Further investigation into differences between housing submarkets and socio-economic groups would also provide a deeper

understanding of how the benefits of green infrastructure are experienced and valued across different populations.

Overall, this research demonstrates that urban green infrastructure plays a meaningful but complex role within metropolitan housing markets. Green spaces can contribute to residential desirability and may generate economic value, but this relationship is shaped by the wider urban context in which these amenities exist. The findings challenge the idea of a universal green premium and instead highlight the importance of understanding greenery as part of a broader system of urban characteristics, accessibility patterns, and social conditions.

In rapidly transforming cities such as London, where pressures of housing demand, land scarcity, and environmental challenges increasingly intersect, recognising this complexity is essential. Green infrastructure should therefore be understood not only as an environmental resource or a market amenity, but as a fundamental component of sustainable urban development capable of improving the quality, resilience, and inclusiveness of future cities.

BIBLIOGRAPHY

Britannica, T. Editors of Encyclopaedia (2024) *Lambeth*. Encyclopaedia Britannica. Available at: https://www.britannica.com/place/Lambeth (Accessed: 22 August 2026).

Census Information Scheme (2022) *2021 census wards housing* [Data set]. London Datastore. Available at: https://data.london.gov.uk/dataset/2021-census-wards-housing (Accessed: 22 August 2026).

Chi, B., Dennett, A., Oléron-Evans, T. and Morphet, R. (2021) *House price per square metre in England and Wales, 1995–2021* [Data collection]. Colchester: UK Data Service. doi: 10.5255/UKDA-SN-855033.

de Vries, S., Verheij, R.A., Groenewegen, P.P. and Spreeuwenberg, P. (2003) 'Natural environments—healthy environments? An exploratory analysis of the relationship between greenspace and health', *Environment and Planning A: Economy and Space*, 35(10), pp. 1717–1731. doi: 10.1068/a35111.

Earnhart, D. (2002) 'Combining revealed and stated data to examine housing decisions using discrete choice analysis', *Journal of Urban Economics*, 51(1), pp. 143–169. doi: 10.1006/juec.2001.2241.

Elliott, R., Motzny, A., Majd, S., Viteri, C.F., Laimer, D., Orlove, B. and Culligan, P. (2019) 'Identifying linkages between urban green infrastructure and ecosystem services using an expert opinion methodology', *Ambio*, 49. doi: 10.1007/s13280-019-01223-9.

Fan, Q., Hansz, J. and Yang, X. (2015) 'The pricing effects of open space amenities', *The Journal of Real Estate Finance and Economics*, 52. doi: 10.1007/s11146-015-9508-1.

Fávero, L. (2011) 'Urban amenities and dwelling house prices in São Paulo, Brazil: a hierarchical modelling approach', *Global Business and Economics Review*, 13, pp. 147–167. doi: 10.1504/GBER.2011.040729.

GLA Intelligence (2015) *Ward profiles and atlas: ward atlas for 2014 boundaries* [Spatial data collection]. London Datastore. Available at: https://data.london.gov.uk/dataset/ward-profiles-and-atlas (Accessed: 22 August 2026). [Note: original source link was broken/incomplete — please verify and update this URL.]

Greater London Authority (2015) *Statistical GIS boundary files for London* [Spatial data collection]. Available at: https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london (Accessed: 22 August 2026).

House of Commons (2017) *Public parks*. Communities and Local Government Committee, Seventh Report of Session 2016–17. London: House of Commons.

Jim, C.Y. and Chen, W. (2006) 'Impacts of urban environmental elements on residential housing prices in Guangzhou (China)', *Landscape and Urban Planning*, 78, pp. 422–434. doi: 10.1016/j.landurbplan.2005.12.003.

Kim, H., Lee, G.E., Lee, J.S. and Choi, Y. (2019) 'Understanding the local impact of urban park plans and park typology on housing price: a case study of the Busan metropolitan region, Korea', *Landscape and Urban Planning*, 184, pp. 1–11. doi: 10.1016/j.landurbplan.2018.12.007.

Kong, F., Yin, H. and Nakagoshi, N. (2007) 'Using GIS and landscape metrics in the hedonic price modelling of the amenity value of urban green space: a case study in Jinan City, China', *Landscape and Urban Planning*, 79(3–4), pp. 240–252.

Kuminoff, N.V. and Pope, J.C. (2014) 'Do "capitalization effects" for public goods reveal the public's willingness to pay?', *International Economic Review*, 55, pp. 1227–1250. Available at: https://www.public.asu.edu/~nkuminof/KP12.pdf(Accessed: 22 August 2026).

London Borough of Tower Hamlets (2013) *Introduction*. Strategic Planning Team, Planning and Building Control Service, Place Directorate. Available at: https://www.towerhamlets.gov.uk/Documents/Planning-and-building-control/Strategic-Planning/Local-Plan/Introduction.pdf (Accessed: 22 August 2026).

Office for National Statistics (2011) *London postcodes* [Data collection]. Available at: https://www.doogal.co.uk/london_postcodes (Accessed: 22 August 2026).

Office for National Statistics (2023) *Jobs and job density, borough* [Data collection]. Available at: https://data.london.gov.uk/dataset/jobs-and-job-density-borough (Accessed: 22 August 2026).

OpenStreetMap (2023) *Greater London* [Spatial data collection]. Available at: https://download.geofabrik.de/europe/great-britain/england/greater-london.html (Accessed: 22 August 2026).

Ordnance Survey (2020) *Map of the London Borough of Tower Hamlets, showing electoral wards*. © Crown copyright and database right. Available at: https://commons.wikimedia.org/w/index.php?curid=15931089 (Accessed: 22 August 2026). Licensed under CC BY-SA 3.0.

Osland, L., Östh, J. and Nordvik, V. (2022) 'House price valuation of environmental amenities: an application of GIS-derived data', *Regional Science Policy & Practice*, 14(4), pp. 939–959. doi: 10.1111/rsp3.12382.

O'Sullivan, A. (2012) 'Why is housing different?', in *Urban Economics*. 8th edn. Boston: McGraw-Hill/Irwin, pp. 367–385.

Panduro, T. and Veie, K. (2013) 'Classification and valuation of urban green spaces – a hedonic house price valuation', *Landscape and Urban Planning*, 120, pp. 119–128.

Ritchie, H. and Roser, M. (2018) *Urbanisation*. Our World in Data. Available at: https://ourworldindata.org/urbanization(Accessed: 22 August 2026).

Rosen, S. (1974) 'Hedonic prices and implicit markets: product differentiation in pure competition', *Journal of Political Economy*, 82, pp. 34–55. doi: 10.1086/260169.

Sassen, S. (1991) *The global city: New York, London, Tokyo*. Princeton, NJ: Princeton University Press.

Schaeffer, Y. and Dissart, J.-C. (2018) 'Natural and environmental amenities: a review of definitions, measures and issues', *Ecological Economics*, 146, pp. 475–496. doi: 10.1016/j.ecolecon.2017.12.001.

Sirmans, S., Macpherson, D. and Zietz, E. (2005) 'The composition of hedonic pricing models', *Journal of Real Estate Literature*, 13(1), pp. 1–44. doi: 10.1080/10835547.2005.12090154.

Su, S., He, S., Sun, C., Zhang, H., Hu, L. and Kang, M. (2021) 'Do landscape amenities impact private housing rental prices? A hierarchical hedonic modelling approach based on semantic and sentimental analysis of online housing advertisements across five Chinese megacities', *Urban Forestry & Urban Greening*, 58, p. 126968. doi: 10.1016/j.ufug.2020.126968.

The World Bank (2023) *Urban development*. Available at: https://www.worldbank.org/en/topic/urbandevelopment/overview (Accessed: 22 August 2026).

Tower Hamlets Council (2007) *Island Gardens Conservation Area*. Available at: https://www.towerhamlets.gov.uk/Documents/Planning-and-building-control/Development-control/Conservation-areas/Island-GardensV1.pdf (Accessed: 22 August 2026).

APPENDIX A - London variables selection R Code

```
import csv

# File paths
input_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo.csv'
output_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo_filtered.csv'

# Column to filter
column_name = 'towncity'
value_to_keep = 'LONDON'

# Open input CSV file with specified encoding
with open(input_file, 'r', newline='', encoding='utf-8') as
file:
    reader = csv.DictReader(file)
    headers = reader.fieldnames

    # Filter rows based on the specified condition
        filtered_rows = [row for row in reader if
row[column_name] == value_to_keep]

# Write the filtered rows to the output CSV file
with open(output_file, 'w', newline='', encoding='utf-8') as
file:
    writer = csv.DictWriter(file, fieldnames=headers)
    writer.writeheader()
    writer.writerows(filtered_rows)

print("Rows filtered and the modified CSV file has been
saved.")
```

APPENDIX B - Excluding Columns R Code

```
import csv

# File paths
input_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo_filtered.csv'
output_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo_filtered_modified.csv'

# Columns to exclude
columns_to_exclude = [
                    "POTENTIAL_ENERGY_RATING",
"POTENTIAL_ENERGY_EFFICIENCY", "INSPECTION_DATE",
            "ENVIRONMENT_IMPACT_POTENTIAL",
"ENERGY_CONSUMPTION_POTENTIAL",
"CO2_EMISSIONS_POTENTIAL",
      "LIGHTING_COST_POTENTIAL", "HEATING_COST_POTENTIAL",
"HOT_WATER_COST_POTENTIAL",
          "ENERGY_TARIFF", "MAINS_GAS_FLAG",
"MAIN_HEATING_CONTROLS", "LOW_ENERGY_LIGHTING",
        "NUMBER_OPEN_FIREPLACES", "HOTWATER_DESCRIPTION",
"HOT_WATER_ENERGY_EFF", "HOT_WATER_ENV_EFF",
          "FLOOR_ENERGY_EFF", "FLOOR_ENV_EFF",
"WINDOWS_DESCRIPTION", "WINDOWS_ENERGY_EFF",
          "WINDOWS_ENV_EFF", "WALLS_DESCRIPTION",
"WALLS_ENERGY_EFF", "WALLS_ENV_EFF",
         "SECONDHEAT_DESCRIPTION", "SHEATING_ENERGY_EFF",
"SHEATING_ENV_EFF", "ROOF_DESCRIPTION",
          "ROOF_ENERGY_EFF", "ROOF_ENV_EFF",
"MAINHEAT_DESCRIPTION", "MAINHEAT_ENERGY_EFF",
          "MAINHEAT_ENV_EFF", "MAINHEATCONT_DESCRIPTION",
"MAINHEATC_ENERGY_EFF", "MAINHEATC_ENV_EFF",
          "LIGHTING_DESCRIPTION", "LIGHTING_ENERGY_EFF",
"LIGHTING_ENV_EFF", "MAIN_FUEL",
          "WIND_TURBINE_COUNT", "HEAT_LOSS_CORRIDOOR",
"UNHEATED_CORRIDOR_LENGTH", "PHOTO_SUPPLY",
     "SOLAR_WATER_HEATING_FLAG", "MECHANICAL_VENTILATION",
"FIXED_LIGHTING_OUTLETS_COUNT",
     "LOW_ENERGY_FIXED_LIGHT_COUNT", "duration", "oldnew",
"propertytype", "towncity", "county",
      "categorytype", "recordstatus", "FLOOR_DESCRIPTION",
"NUMBER_HEATED_ROOMS",
          "GLAZED_TYPE", "FLAT_TOP_STOREY",
"HOT_WATER_COST_CURRENT", "HEAT_COST_CURRENT",
                          "LIGHT_COST_CURRENT",
"CO2_EMISS_CURR_PER_FLOOR_AREA", "CO2_EMISSIONS_CURRENT",
       "ENVIRONMENT_IMPACT_CURRENT", "ENERGY_CONSUMPTION",
"CURRENT_ENERGY_EFFICIENCY"
]
```

```
# Open input CSV file with specified encoding
with open(input_file, 'r', newline='', encoding='utf-8')
as file:
    reader = csv.DictReader(file)
    headers = reader.fieldnames

    # Determine the indices of columns to keep
         indices_to_keep = [i for i, header in
enumerate(headers) if header not in columns_to_exclude]

    # Filter and keep only the desired columns
      filtered_rows = [{headers[i]: row[headers[i]] for i
in indices_to_keep} for row in reader]

# Write the filtered rows to the output CSV file
with open(output_file, 'w', newline='', encoding='utf-8')
as file:
     writer = csv.DictWriter(file, fieldnames=[headers[i]
for i in indices_to_keep])
    writer.writeheader()
    writer.writerows(filtered_rows)

print("Columns removed, and the final modified CSV file
has been saved.")
```

APPENDIX C *- Deleting Insignificant Values From Floor Area R Code*

```
import csv

# File paths
input_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo_filtered_modified.csv'
output_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo_filtered_modified_2.csv'

# Column to filter
column_name = 'TOTAL_FLOOR_AREA'
value_to_exclude = '0'

# Open input CSV file with specified encoding
with open(input_file, 'r', newline='', encoding='utf-8')
as file:
    reader = csv.DictReader(file)
    headers = reader.fieldnames

    # Filter rows based on the specified condition
          filtered_rows = [row for row in reader if
row[column_name] != value_to_exclude]

# Write the filtered rows to the output CSV file
with open(output_file, 'w', newline='', encoding='utf-8')
as file:
    writer = csv.DictWriter(file, fieldnames=headers)
    writer.writeheader()
    writer.writerows(filtered_rows)

print("Rows filtered, and the modified CSV file has been
saved.")
```

```
import csv

# File paths
input_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\allinfo_filtered_modified_2.csv'
output_file = r'C:\Users\giova\OneDrive\Documents\Área de
Trabalho\allinfo\pre_final.csv'

# Column to filter
column_name = 'CONSTRUCTION_AGE_BAND'
values_to_exclude = ['INVALID!', 'NO DATA!', '']

# Open input CSV file with specified encoding
with open(input_file, 'r', newline='', encoding='utf-8')
as file:
    reader = csv.DictReader(file)
    headers = reader.fieldnames

    # Filter rows based on the specified condition
          filtered_rows = [row for row in reader if
row[column_name] not in values_to_exclude]

# Write the filtered rows to the output CSV file
with open(output_file, 'w', newline='', encoding='utf-8')
as file:
    writer = csv.DictWriter(file, fieldnames=headers)
    writer.writeheader()
    writer.writerows(filtered_rows)

print("Rows filtered, and the modified CSV file has been
saved.")
```

APPENDIX E - Deleting Rows Outside City Center R Code

```
import csv

# File paths
input_file = r'C:\Users\giova\OneDrive\Documents\Tese\allinfo\1_final.csv'
output_file = r'C:\Users\giova\OneDrive\Documents\Tese\allinfo\1_final_filtered.csv'

# Column to filter
column_name = 'district'
values_to_keep = ["CAMDEN", "ISLINGTON", "HACKNEY", "TOWER HAMLETS", "SOUTHWARK", "LAMBETH", "CITY OF WESTMINSTER", "WESTMINSTER" , "KENSINGTON AND CHELSEA", "CITY OF LONDON"]

# Open input CSV file with specified encoding
with open(input_file, 'r', newline='', encoding='utf-8') as file:
    reader = csv.DictReader(file)
    headers = reader.fieldnames

    # Filter rows based on the specified condition
        filtered_rows = [row for row in reader if row[column_name] in values_to_keep]

# Write the filtered rows to the output CSV file
with open(output_file, 'w', newline='', encoding='utf-8') as file:
    writer = csv.DictWriter(file, fieldnames=headers)
    writer.writeheader()
    writer.writerows(filtered_rows)

print("Rows filtered, and the modified CSV file has been saved.")
```