

Group Name's Group Project

Declaration of Authorship

We, [python high Group], confirm that the work presented in this assessment is our own. Where information has been derived from other sources, we confirm that this has been indicated in the work. Where a Large Language Model such as ChatGPT has been used we confirm that we have made its contribution to the final submission clear.

Date: 19/December/2023

Student Numbers: 23237413, 23212209, 23075581, 23092081, 23041170

Brief Group Reflection

Went well:

- 1.The selected analysis method effectively analyzes the spatial differentiation characteristics of the relationship between Airbnb and deprivation index.
 - 2.The findings of the study are well suited to complement the policy challenge
-
- 1.The large amount of data in Airbnb's data set poses a challenge to the preprocess of data
 - 2.Choosing the right parameters and specifications for the GWR model and getting the model to run successfully in JupyterLab was a challenge.

Priorities for Feedback

Are there any areas on which you would appreciate more detailed feedback if we're able to offer it?

We would like to get feedback on methodology, especially the use of GWR. Also, we would like the feedback on the broader implications of findings and suggestions for policy recommendations to help us enhance the practical relevance.

Response to Questions

This study explores the spatial distribution pattern of Airbnb using cluster analysis and constructs its relationship with Deprivation using the GWR model, which aims to provide evidence for optimising Airbnb's policy and management mode. The materials needed for the study have been saved in the following link, to avoid errors, please download the data directly when reproducing the code.

https://github.com/songzimen/FSDS_Deprivation_Airbnb_GWR.git

```
/tmp/ipykernel_2870/3569245906.py:3: DtypeWarning:
```

```
Columns (17) have mixed types. Specify dtype option on import or set low_memory=False.
```

1. Who collected the data?

...

The dataset for the study consists of two parts. The Airbnb data comes from **Inside Airbnb**, while the Deprivation data is from **the Department for Levelling Up, Housing & Communities**, which is part of the UK government.

(2 points; Answer due Week 7)

...

2. Why did they collect it?

...

Inside Airbnb collects publicly available data from the Airbnb website to provide data and advocacy about Airbnb's impact on residential communities. Their aim is to enable residents to better understand and decide how to manage the affairs of renting out their homes to tourists and thereby play a greater role in the short-term letting market.

The Department for Levelling Up, Housing & Communities (2019) regards the Deprivation Indices as a useful tool to measure the relative distribution of deprivation within a small area. The dataset is designed to help government and the public better understand socio-economic disparities and optimise decision-making and resource allocation accordingly.

(4 points; Answer due Week 7)

...

3. How was the data collected?

::

Inside Airbnb, created by tech expert Murray Cox and supported by many urban research specialists, uses a suite of open-source technologies, including D3, Bootstrap, Python, PostgreSQL, and Google Fonts for compiling and analyzing publicly available data from the Airbnb platform.

The deprivation data, as detailed in the “English Indices of Deprivation 2019” document, is gathered using a comprehensive approach. This process involves the collection and integration of various socio-economic indicators from multiple government data sources. These indicators cover diverse aspects such as income, employment, education, health, crime, and barriers to housing and services. The data is then meticulously compiled to create a detailed picture of socio-economic conditions across geographic areas. This method ensures an accurate and multi-faceted view of deprivation at a localized level.

(5 points; Answer due Week 8)

:::

4. How does the method of collection impact the completeness and/or accuracy of its representation of the process it seeks to study, and what wider issues does this raise?

4.1 Data Source and Timing

The Deprivation Indices are based on comprehensive statistical data from various government departments, each with different data collection frequencies and scales. This results in extended intervals between updates (2007, 2010, 2015, 2019). Any changes in data collection frameworks by these departments can significantly impact the indices. Particularly, the lack of updates during pivotal events like the 2020-2022 global pandemic might limit the dataset's current relevance to urban development. In contrast, Airbnb's data is more dynamic and frequently updated, capturing immediate changes despite being sourced solely from an online platform. These differences in update frequency and source create disparities in temporal scales between the two datasets, posing challenges in accurately linking Airbnb's influence on community deprivation.

4.2 Data Quality and Reliability

The Indices of Deprivation benefit from stringent quality assurance processes, including comprehensive checks and validation. In contrast, the collection methods for Airbnb data, such as web scraping or API, might adhere to different standards of quality control. This variation in quality assurance practices could affect the reliability of the Airbnb data, especially when it's compared with the well-validated Deprivation Indices. Such differences in data quality and control methods highlight potential challenges in assessing the datasets' comparability and reliability.

4.3 Wider Issues

- **Dynamic Relationship:** The link between Airbnb's presence and community deprivation is constantly evolving, reflecting changes in Airbnb's market status. This fluidity must be carefully considered during analysis.
- **Timeframe Discrepancies:** The different update intervals of the Deprivation Indices and Airbnb data can lead to challenges in accurately depicting their current relationship. Airbnb's real-time updates, influenced by timely and specific factors, add another layer of complexity.
- **Data Quality and Ethical Implications:** The quality and reliability of Airbnb's data, potentially less controlled than the Deprivation Indices, introduce ethical concerns. Variations in data accuracy or potential biases could significantly affect the credibility of the research findings.

(11 points; Answer due Week 9)

5. What ethical considerations does the use of this data raise?

5.1 Privacy and Data Protection

Handling both the Indices of Deprivation and Airbnb data demands strict privacy considerations. Sensitive details, like host names or personal photos, should be carefully excluded or anonymized. Any unavoidable use of such private data requires stringent protective measures and necessary permissions to ensure compliance with data protection laws.

5.2 Potential Stigmatization

The combined use of deprivation indices and Airbnb data carries the risk of inadvertently stigmatizing certain areas as "deprived" or "underprivileged." This could influence public perception and economic activities in these regions, affecting local residents and potentially skewing Airbnb investment and user preferences.

5.3 Bias and Accuracy in Data Representation

The varying data collection methods and update frequencies between the two datasets might lead to biases. Ethical research practices necessitate transparent acknowledgment of these potential biases and limitations of the study.

5.4 Responsibility in Data Interpretation

Researchers must responsibly interpret and present their findings, considering the dynamic nature of Airbnb's market influence and the socio-economic factors reflected in the deprivation data.

(18 points; Answer due ?var:assess.group-date)

6. With reference to the data (i.e. using numbers, figures, maps, and descriptive statistics), what does an analysis of Hosts and Listing types suggest about the nature of Airbnb lets in London?

The provided image below presents an analysis of Airbnb's room types in London, focusing on the quantity and average price, alongside a detailed view of the top 10 hosts' offerings. Here are the refined key insights.

- “Entire home/apt” is the predominant listing category and, together with “Hotel room”, garners a higher price point than other accommodation types.
- The majority of the most active hosts in London list “Entire home/apt” options, suggesting a prevalent strategy of renting out entire properties.
- There’s a notable diversity in listing prices across room types and hosts, implying that prices are set variably based on location, property quality, and the services provided.

(15 points; Answer due ?var:assess.group-date)

:::

/tmp/ipykernel_2870/2587134144.py:15: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

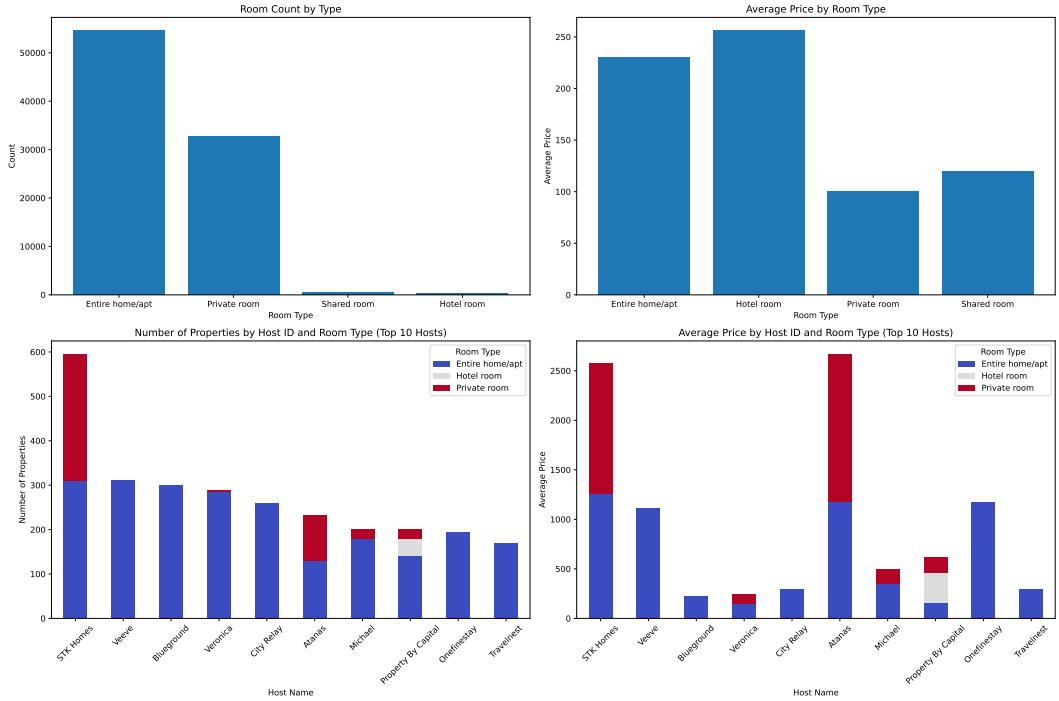
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#inplace-mutation

/tmp/ipykernel_2870/2587134144.py:21: FutureWarning:

The default of observed=False is deprecated and will be changed to True in a future version of pandas.

/tmp/ipykernel_2870/2587134144.py:24: FutureWarning:

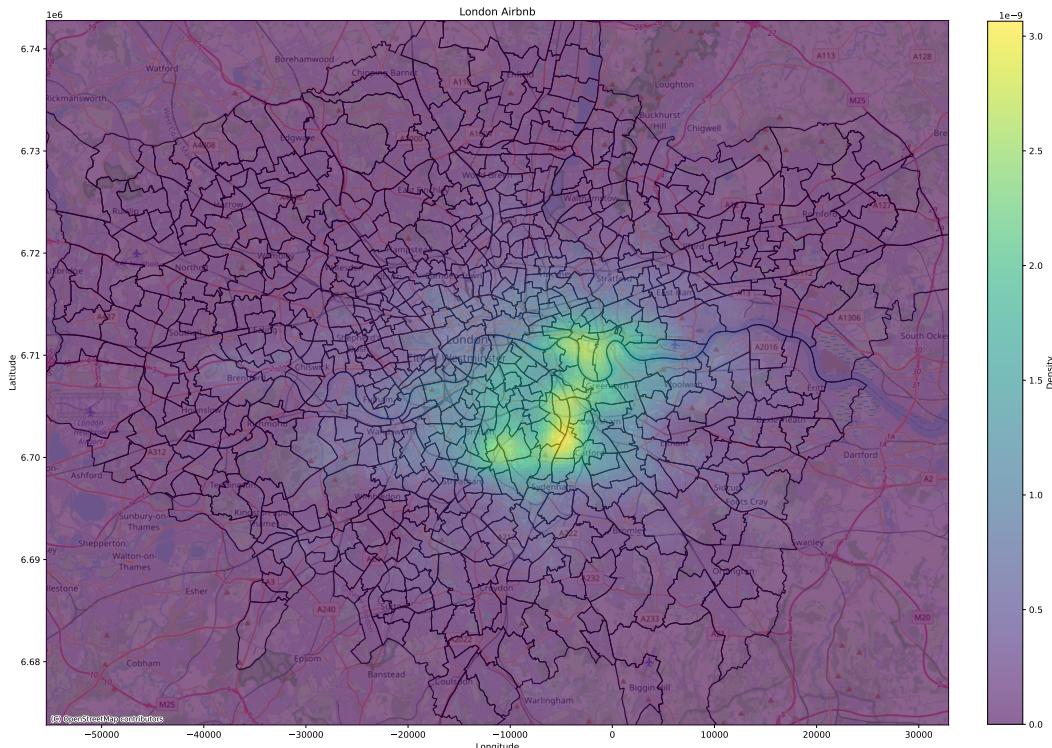
The default of observed=False is deprecated and will be changed to True in a future version of pandas.



The heatmap below illustrates the concentration of Airbnb listings across London, showing a hotspot in the central region that gradually fades outwards. The densest area, highlighted in bright yellow, indicates a cluster of listings located to the east and south of central London, an area known for attractions like Tower Bridge, Canary Wharf, and Greenwich Observatory. This aligns with the bar chart showing a high volume of 'Entire home/apt' listings, suggesting hosts are meeting the strong demand for accommodations in these desirable locations, reflective of London's varied socio-economic landscape.

```
/opt/conda/lib/python3.11/site-packages/pyproj/crs/crs.py:141: FutureWarning:
```

```
'+init=<authority>:<code>' syntax is deprecated. '<authority>:<code>' is the preferred initia
```



7. Drawing on your previous answers, and supporting your response with evidence (e.g. figures, maps, and statistical analysis/models), how could this data set be used to inform the regulation of Short-Term Lets (STL) in London?

7.1 Introduction: The Rise of the Sharing Economy and the Urgency for Regulation

Amidst the sharing economy's surge, London's STL market, propelled by platforms like Airbnb, has dramatically influenced urban housing and community dynamics, prompting calls for regulation to safeguard housing market stability and community welfare. This study harnesses spatial auto-correlation to assess Airbnb distribution patterns and geographic weighted regression (GWR) to examine the influence of deprivation sub-indicators on these prices. By discerning the geographical distribution of Airbnb and the differential impact of deprivation sub-indicators, the research provides deeper insights for policymakers. The integration of these analyses aims to inform STL regulatory strategies, ensuring market growth aligns with community interests and broad socioeconomic context.

(45 points; Answer due ?var:assess.group-date)

7.2 Spatial Patterns Unveiled: Significant Spatial Segregation in Airbnb Listings

Are there spatial distributional disparities in the number and price of Airbnb's in London? If exist, what characteristics do they reflect? To answer this question, the study first conducted a global spatial auto-correlation analysis and calculates the Moran's indices for the two types of spatial data to be 0.76 and 0.43. These

results confirm that there are statistically significant spatial distribution differences in both the number of Airbnb listings and the average price of properties across different areas of London.

Global Moran's I: 0.7554256506929731

The z-test value under random distribution hypothesis: 32.283209075386196

The p-value of the z-test under random distribution hypothesis: 1.2035930392272867e-228

The z-test value under normal distribution hypothesis: 32.00800184002635

The p-value of the Z test under normal distribution hypothesis: 8.438495568419313e-225

Global Moran's I: 0.4251755584591877

The z-test value under random distribution hypothesis: 18.23100860112638

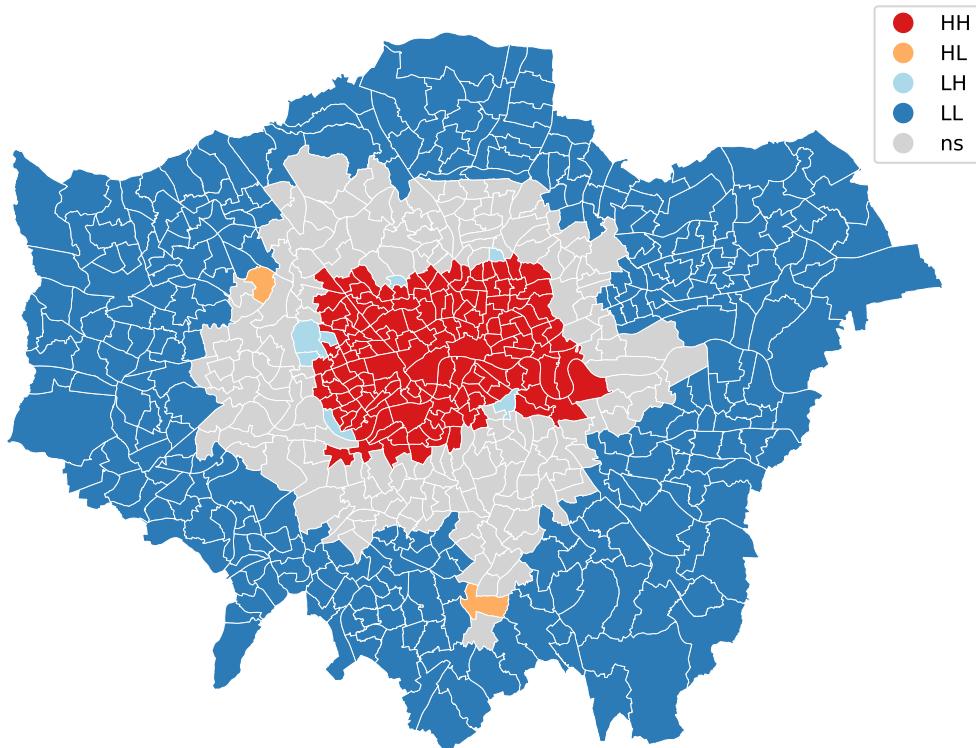
The p-value of the z-test under random distribution hypothesis: 2.9286978716824874e-74

The z-test value under normal distribution hypothesis: 18.044657603586156

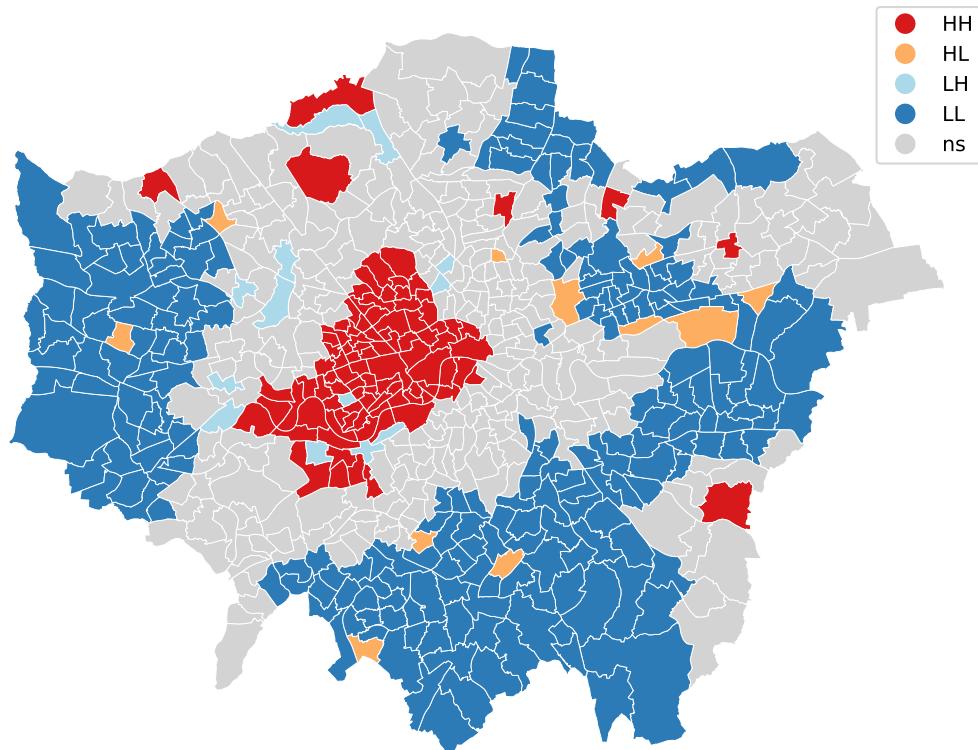
The p-value of the Z test under normal distribution hypothesis: 8.690138055356104e-73

To delve deeper into the characteristics of the spatial distribution of Airbnb listings and their average pricing, the study performed a local Moran's I analysis. This highlights the clustering of high and low values in different areas (see figure below).

LISA Cluster Map for Count_



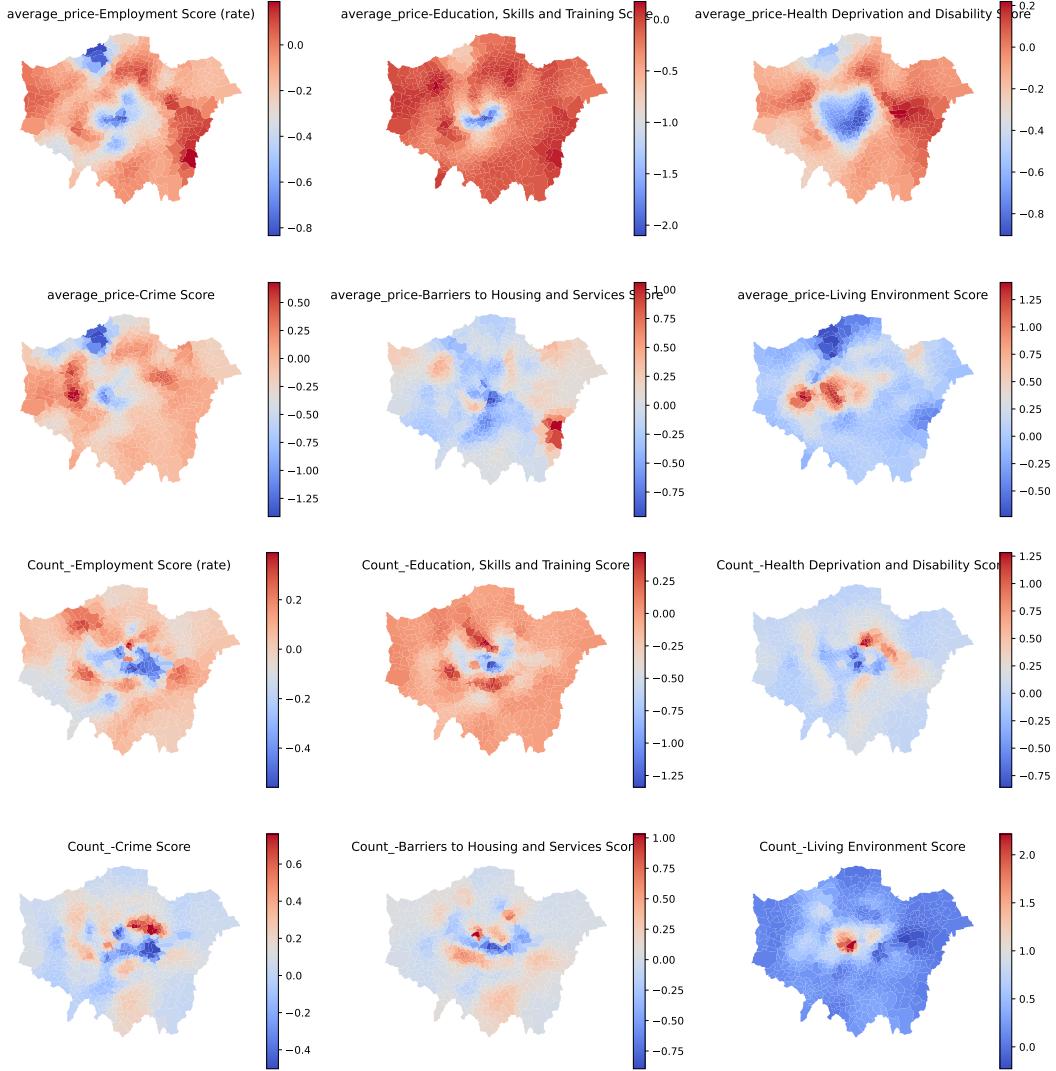
LISA Cluster Map for average_price



The cluster maps show that the distribution of Airbnb listings quantity is significantly clustered in inner and dispersed in outer London areas. The average price clusters are more scattered, with hotspots primarily in the central southwestern part of London and cold spots in the west, north, and southeast. The abundance of tourist resources and infrastructure in Inner London may contribute to a noticeable clustering of Airbnb's quantity and pricing in these areas. There is a clear demarcation between areas of high and low values, indicating significant differences in economic conditions and accommodation demand between these regions. Airbnb's average pricing is more susceptible to the degree of poverty in an area (Gutiérrez, García-Palomares, Romanillos, & Salas-Olmedo, 2017). We will now proceed to further investigate the relationship between seven poverty indicators and average pricing using GWR.

7.3 Intersecting Correlations: Airbnb Pricing and Deprivation

Given that Airbnb listing prices have an irregularly spatial distribution featuring clustering, we would like to explore the link between prices and the indicators of Deprivation. Imagine whether the prices of Airbnb listings are affected by the income and education level of the residents in the neighbourhood where they are located?



Our Geographically Weighted Regression (GWR) analysis reveals a distinct pattern in central London boroughs like the City of London, Westminster, Kensington, Chelsea, and Hammersmith and Fulham. Here, Airbnb prices negatively correlate with indicators such as employment, health, education, and crime rates. This trend suggests that areas with higher deprivation scores see lower Airbnb prices, affected by factors like crime rates and limited infrastructure. Despite their central locations, these areas are less appealing to renters.

The trend shifts as we move to the suburbs. The negative correlation between deprivation scores and Airbnb pricing weakens in areas like Barnet and Harrow. These suburban regions show a different dynamic, where high housing supply and lower demand drive prices down. A notable deviation is observed in living environment scores. In central areas, the aged infrastructure does not deter high Airbnb prices, likely due to the allure of prime locations and tourist attractions. In contrast, suburban Airbnb listings show a negative correlation, indicating a preference for newer, well-maintained properties in these less crowded areas.

7.4 Conclusion\xWhat can policy do?

Our analysis indicates that in Inner London, Airbnb prices are negatively impacted by high poverty levels, particularly in areas with challenges in employment, educa-

tion, health, and disability. This situation calls for a customized approach to Airbnb regulation in these neighborhoods.

- Inner London Measures:

For Inner London, we propose two key strategies to support diverse groups and enhance the Airbnb market. Firstly, we recommend economic incentives for special groups including individuals with disabilities, health issues, the unemployed. By offering short-term rental discounts to these groups, landlords would receive subsidies, fostering social equity and making housing more affordable. Secondly, we advocate for supportive partnerships with government rental programs (KeyNest, 2019). These collaborations can provide security and financial benefits to landlords in disadvantaged areas, creating a more stable and equitable rental market in Inner London.

- Outer London Strategy:

In Outer London, our focus shifts to elevating the quality of Airbnb listings to boost market competitiveness. The strategy here emphasizes prioritizing high-quality listings. Implementing a quality assessment system will help regulate and elevate the standard of accommodations, ensuring only the best are available for rent. This measure aims to attract guests seeking high-standard accommodations and positions Outer London's Airbnb market as a provider of premium lodging experiences. By concentrating on quality over quantity,

Sustainable Authorship Tools

Your QMD file should automatically download your BibTeX file. We will then re-run the QMD file to generate the output successfully.

Written in Markdown and generated from [Quarto](#). Fonts used: [Spectral](#) (mainfont), [Roboto](#) (sansfont) and [JetBrains Mono](#) (monofont).

References

'bio.bib' nocite: " ...option]