

Impact of Airbnb on local area rental markets

Yellow Submarine's Group Project

Declaration of Authorship

We, [Yellow Submarine], pledge our honour 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: 17/12/2024

Student Numbers:

s.lin.24@ucl.ac.uk Shijie Lin
wenkai.song.24@ucl.ac.uk Wenkai Song
yifan.feng.23@ucl.ac.uk Yifan Feng
hanbing.xuan.24@ucl.ac.uk Hanbing Xuan
bowen.wu.24@ucl.ac.uk Bowen Wu

Brief Group Reflection

What Went Well	What Was Challenging
analysis	the rendering part combine all the work together

Priorities for Feedback

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

Analysis process.

Response to Questions

All the datasets, shapefiles, bibfile could be found on
https://github.com/wbwhaha/FSDS_Group_Assignment.

The thirteenth code cell may take more than a minute to read the data, depending on the quality of the network.

1. Who collected the InsideAirbnb data?

Inside Airbnb is a collective of residents, activists, and allies from a range of organizations worldwide ('Inside airbnb' (n.d.)).

2. Why did they collect the InsideAirbnb data?

As discussed on 'Inside airbnb' (n.d.), the data collected by Inside Airbnb is driven by their mission to support residents and activists organizing to shield their communities from potential negative impacts of short-term rentals. The collective believes in empowering communities through shared resources and by fostering collaboration. Their network acts as an inclusive and safe space for residents, activists, and allies to connect, learn from one another, and organize collective action. It aims to amplify efforts and pool resources to mitigate the challenges posed by short-term rental platforms.

3. How did they collect it?

The Inside Airbnb website leverages a range of open-source technologies and modern services for data presentation and hosting. Key technologies include **D3.js** for dynamic data visualization, **Bootstrap** for creating responsive, mobile-friendly layouts, **Python** for data scraping and analysis, such as extracting publicly available housing information, prices, reviews, and other dynamic content from the Airbnb website, **PostgreSQL** for managing structured datasets, and **Google Fonts** for enhancing design aesthetics. Maps on the site are designed using **Mapbox** with data sourced from **OpenStreetMap** and are hosted through Mapbox's robust API services. The website itself is hosted via **Amazon S3**, ensuring fast, secure, and reliable access.

4. How does the method of collection (Q3) impact the completeness and/or accuracy of the InsideAirbnb data? How well does it represent the process it seeks to study, and what wider issues does this raise?

Impact of Data Collection Method on Completeness/Accuracy:

Data Timeliness:

Inside Airbnb data reflects a specific point in time, missing real-time updates and potentially creating discrepancies with the current market.

Selection Bias:

The scraper may exclude inactive or temporarily delisted properties, reducing data

comprehensiveness and reliability, which was mentioned in Barron, Kung and Prosperio (2018).

Data Integrity:

Website structure changes or restrictions can cause data loss, affecting quality and credibility.

Representation of the Process Studied:

Real-Time Limitations:

As showed in Gurran and Phibbs (2017), while valuable for analyzing market impacts, the data lacks real-time updates, hindering dynamic trend analysis.

Theoretical vs. Reality:

The dataset aids economic research but may underrepresent areas or property types due to selection bias, limiting practical applications.

Wider Issues Raised:

Privacy and Data Ethics:

Scraping raises concerns about user privacy and ethical use of personal data.

Policy Impact:

The data's accuracy influences policymaking on short-term rental regulation.

Community Response:

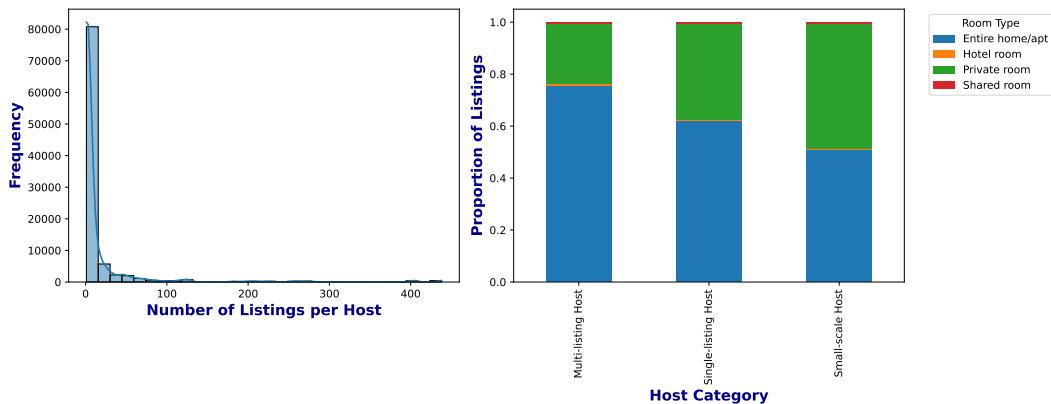
Data presentation may shape public perception, either encouraging rentals or amplifying community opposition.

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

When using Inside Airbnb data for research, it is necessary to balance its research value with potential ethical risks. The acquisition of data is often done without the consent of the landlord or tenant, which may violate privacy and the terms of service of the Airbnb platform, leading to legal and ethical disputes, which was noted by Spier (2024). In addition, The results of the Airbnb platform are typically used to support city policies such as short-term rental restrictions. However, as discussed by Prentice and Pawlicz (2024), in the policy-making process, the crawled data is only a snapshot of a specific time point, which may have bias. If the interests of small landlords or the needs of tenants are not fully considered, the analysis conclusions may not be comprehensive enough or mislead policy makers, resulting in unfair impacts on the landlord or tenant groups.

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

Fig_0_a: Distribution of Host Listing Counts Fig_0_b: Room Type Distribution by Host Category



The division standard of host could be found in Wachsmuth and Weisler (2018).

Host Distribution: Most hosts have a small number of listings, indicating that the market is dominated by individual or small-scale hosts rather than large commercial operators. According to (Meris (2023)), the majority of hosts have fewer than five listings. This suggests that many people are using Airbnb as a supplementary income source rather than a full-time business.

Property Types: Entire homes/apartments and private rooms are the most common types of properties listed. Multi-listing hosts tend to list entire homes/apartments more frequently, while single-listing hosts and small-scale hosts list private rooms more often. This pattern is also reflected in the data provided by Smarhost, which shows a high demand for entire homes and private rooms (*The insider's guide to airbnb occupancy rates in london from smarhost* (no date)).

Commercial vs. Personal Use: The prevalence of entire homes/apartments and private rooms suggests that many listings are for personal use rather than commercial use. (*London, airbnb market statistics & data, the united kingdom* (no date)) pointed out that this observation aligns with Airbtcs Market Statistics, which indicates that individual hosts predominantly list entire homes and private rooms.

Overall, the analysis suggests that Airbnb in London is characterized by a large number of individual hosts offering personal accommodations, with entire homes and private rooms being the most common types of listings.

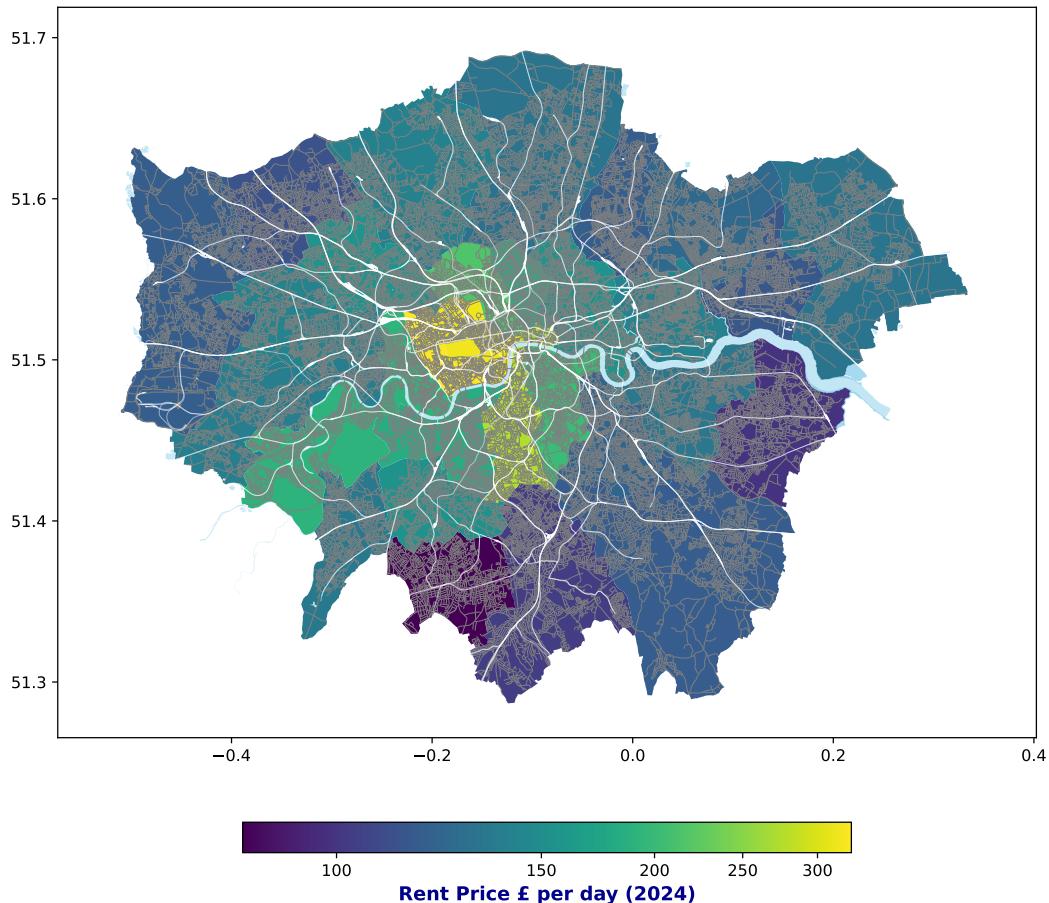
7. Drawing on your previous answers, and supporting your response with evidence (e.g. figures, maps, EDA/ESDA, and simple statistical analysis/models drawing on experience from, e.g., CASA0007), how could the InsideAirbnb data set be used to inform the regulation of Short-Term Lets (STL) in London?

0.BEGINNING

The following analysis including five steps aims to unpack the nature of Airbnb lettings in London by progressively exploring key data dimensions. Starting with a broad geographical perspective, we then narrow our focus to pricing, market competitiveness, commercial hosting patterns, and demand signals, each step building on the insights gleaned from the previous one.

1. Geographic Distribution

Fig_1: Rent Price in London Boroughs (2024)



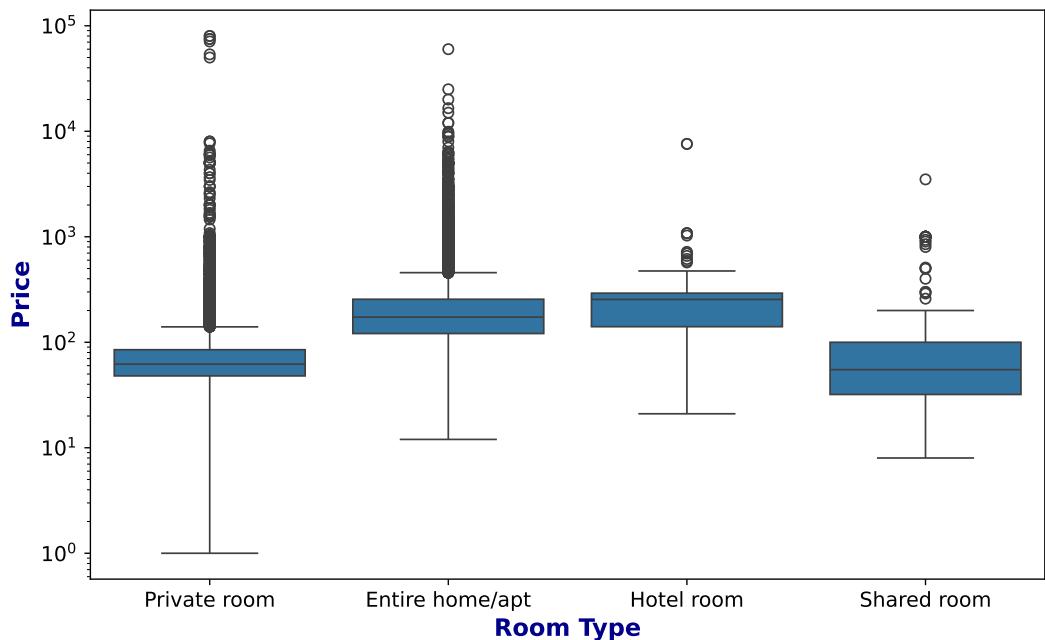
1.1 Price Distribution by boroughs

This map visualizes the spatial distribution of average Airbnb prices across London boroughs, helping identify price variations by area. Yellow areas represent the highest prices, concentrated in central London boroughs like **Westminster** and **Kensington**. Purple and blue areas indicate lower prices, mostly in outer London boroughs. Central London has significantly higher Airbnb listing prices due to higher demand and property values, while prices drop in the outer areas.

2. Listing Price and Type Analysis:

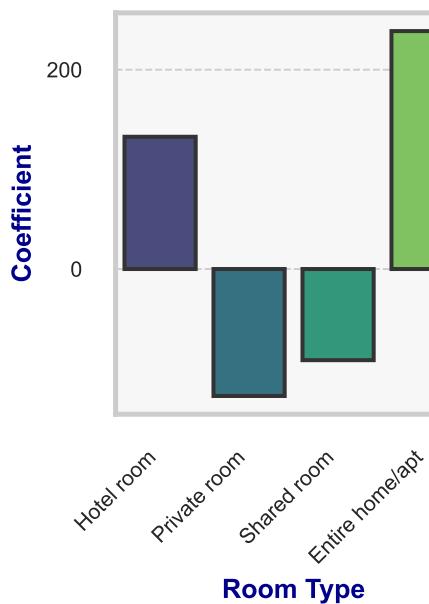
Next, investigate how room types (e.g., entire apartments vs. private rooms) relate to pricing across neighborhoods. Understanding price variations and property types helps clarify whether certain areas and room types cater to more lucrative, tourist-focused niches.

Fig_2_a:Price Distribution by Room Type



2.1 Price Distribution by Room Type The boxplot compares the price distributions of different room types, highlighting medians, spreads, and outliers. Entire home/apt has the highest median price and the widest spread, with many outliers indicating luxury properties. Hotel room prices are concentrated and relatively high but less variable. Private room and Shared room show significantly lower prices and narrower spreads. Entire home/apt and Hotel room are the more expensive options, while Private room and Shared room are the most affordable choices.

Fig_2_b:Impact of Room Type on Price



2.2 Impact of Room Type on Price (Linear Regression Coefficients)

The bar chart visualizes the coefficients from the linear regression model, showing how each room type impacts the price relative to others. Entire home/apt has the largest positive impact on price (239), making it the most expensive type. Hotel room shows a slight positive effect (133). Private room and Shared room both have negative coefficients (-127 and -91), indicating they reduce the price compared to other types. Hotel room significantly increases prices, while Private room and Shared room lead to much lower prices.

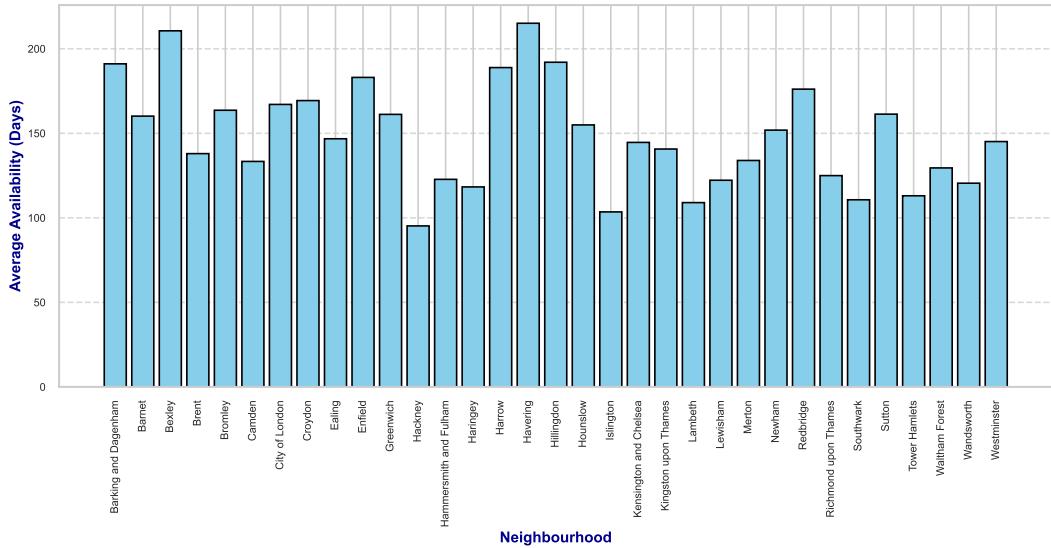
2.3 Summary

Airbnb prices are highest in central London and decrease towards the outskirts. Room type significantly influences price: Hotel room listings have the greatest positive impact, followed by Entire home/apt. In contrast, Private room and Shared room are associated with much lower prices. The increase in individual landlords, who primarily manage shared and private rooms, tends to lower rental prices. Conversely, commercial landlords, who typically oversee hotel rooms and entire apartments, tend to drive prices higher. This distinction reflects a broader trend in the rental market where private landlords contribute to more affordable housing options, such as private room and shared room, while commercial landlords are associated with higher rental costs.

3. Short-Term vs. Long-Term Market Competition:

With a grasp of spatial and price dynamics, add an availability dimension. Determine if listings are predominantly short-term oriented (high availability and short minimum stays), suggesting competition with long-term rental markets.

Fig_3:Average Availability by Neighbourhood



3.1 Analysis of Neighborhood Availability Trends The neighborhoods with the highest average availability are Barnet and Haringey, both exceeding 200 days, while those with the lowest average availability are Camden and Hounslow, both below 100 days. Most neighborhoods have an average availability between 100 and 200 days, showing a general trend. There is significant variability in average availability across different neighborhoods, indicating that some areas have much higher or lower availability than others.

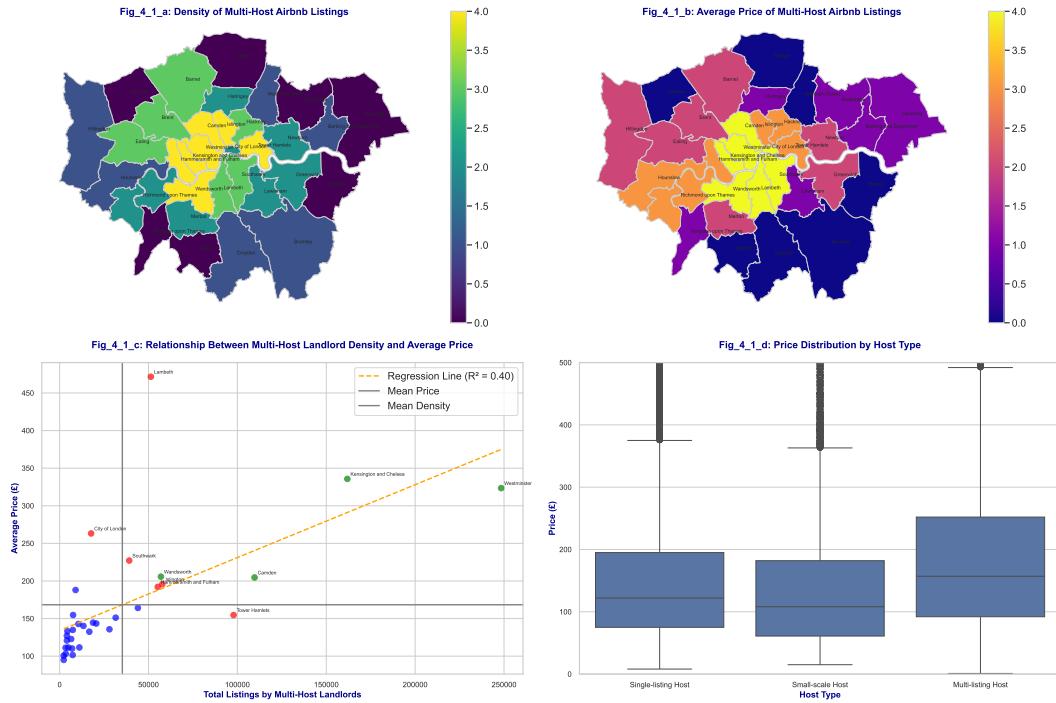
3.2 Impact of High Active Housing Ratio and Total Listings on Long-Term Housing Supply

R-squared (coefficient of determination): R-squared: 0.21219111934091328 which seems good.

For high active housing ratio (high_active_ratio), it has a positive impact on long-term housing supply, that is, when the proportion of high active short-term rental housing increases, the supply of long-term rental housing may also increase.

4. Impact of Commercial Hosts:

After mapping out where and how short-term oriented these listings are, analyze the role of hosts with multiple listings. Identifying commercial operators provides insights into whether Airbnb activity resembles professional hospitality services or remains closer to a home-sharing ethos. In this part, Commercial Host has another name, Multi-listing Host.



4.1 Spatial autocorrelation distribution of data

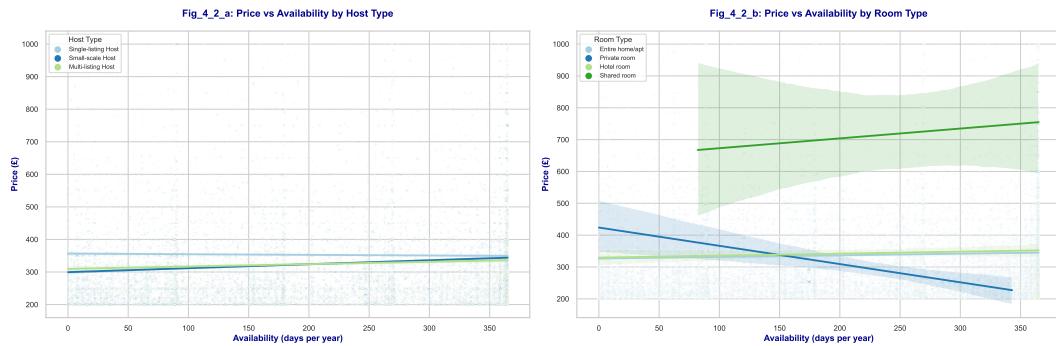
Fig_4_1_a and Fig_4_1_b depict the density and average price of multi host Airbnb listings in London. The central administrative region displays high density and prices, reflecting tourism demand and high property value. The density of peripheral administrative areas has decreased, and prices are generally lower. Although location has a significant impact on density and price, other factors such as property quality, amenities, and competition also play a role.

4.2 Regression analysis

Regression analysis shows a moderate positive correlation ($R^2=0.40$) between multi landlord density and average Airbnb prices. However, important outliers, such as Lambeth's high price despite its low density, highlight the influence of other factors. Administrative regions are classified by price and density: green dots indicate high density and high price, indicating that multi landlord markets have significant influence; Red dots indicate high density or high price; The blue dot represents low density and low price as a baseline. This indicates that the influence of multiple landlords varies by administrative district.

4.3 Price distribution of different landlord types (box plot)

From the median price perspective, the price distribution of multi property landlords is higher and wider, which may indicate that multi property landlords offer more high-end or high priced properties. Single property landlords and small-scale landlords offer lower prices, which may be more suitable for budget travelers.



4.4 Further Analysis

Fig 4_2_a: Relationship between Landlord Type and Price/Length of Stay

1. Multi-landlord property prices increase slightly with rental days, but this effect is weak, suggesting pricing is not heavily influenced by short-term demand.
2. Rental days for all landlord types cluster within 0–200 days, especially multi landlords, indicating a preference for short-term rentals.

Fig 4_2_b: Relationship between Room Type and Price/Rental Days

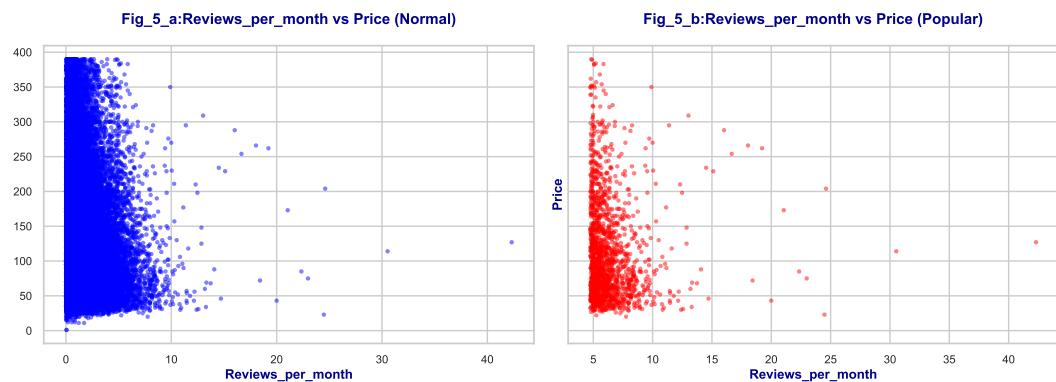
1. Entire homes/apartments show a positive price-length correlation, suited for longer stays.
2. Hotel rooms also increase slightly in price as stays lengthen.
3. Shared and private rooms show a negative price-length relationship, appealing to budget-conscious or temporary tenants.

4.5 Summary

Pricing correlates with central London location and host type. Multi landlords offer broader price ranges but show weak price-term correlations, suggesting stable pricing despite demand fluctuations. Entire properties' prices rise with stay length, while shared rooms drop. These findings highlight the complex dynamics of landlord types, pricing strategies, and market segmentation in London's short-term rental scene.

5. Reviews and Market Demand

Reviews serve as a proxy for demand, with high-review areas reflecting strong tourist interest. Linking demand signals with spatial patterns, pricing, availability, and host types provides a comprehensive view of Airbnb's character in the city.



5.1 Relations Between Price and Popularity

In the “Normal” dataset (low review activity), prices are broad across a mostly low-review landscape, but listings with more than 10 reviews are rare. No clear link between price and review count emerges.

In the “Popular” dataset (high review activity), listings are more scattered, and prices vary widely regardless of review count. No clear upward or downward trend in price aligns with increased reviews.

5.2 Methodology

Using review activity (Reviews_per_month) as a marker, both “Popular” and “Normal” datasets show no evident linear relationship between price and review count. Price remains broadly distributed with no clear correlation to review frequency.

5.3 Findings

Regional Variations:

Popular listings are unevenly distributed across London, with top areas hosting many more listings than lower-ranked regions.

Tourism Hotspots:

High concentrations of popular listings appear in central, tourist-heavy areas (e.g., Westminster, Camden, Tower Hamlets). Such districts attract more visitors and thus have more frequently reviewed properties.

5.4 Summary

Popular listings cluster in central, well-known tourist boroughs, underscoring a strong link between location and listing popularity.

6. Reflection

The InsideAirbnb dataset provides granular insights into London’s STL market and can guide evidence-based regulation, as seen in cities like New York and Barcelona Wachsmuth and Weisler (2018). Policymakers can identify neighborhoods with high STL density, such as Westminster or Kensington, where Airbnb listings cluster around tourist attractions, potentially reducing long-term housing supply.

This data helps distinguish between casual hosts who let rooms occasionally and commercial operators managing multiple properties. Cities like Amsterdam have introduced tiered regulation, placing tighter controls on commercial hosts Nieuwland and Melik (2020). London authorities could use InsideAirbnb data to enforce existing measures—such as the 90-day rule—more effectively. By correlating availability with long-term housing data, they can determine if excessive STR activity is displacing local residents, justifying caps on annual rental days or introducing registration systems.

Additionally, price distributions reveal whether entire homes in central areas command excessive rates, informing price caps or additional taxes. The dataset’s room-type comparisons can help promote affordable options while curbing excessive commercial listings. Beyond reactive measures, policymakers can track the impact of new rules over time, adjusting strategies as needed.

By using InsideAirbnb’s detailed listing-level, spatial, and temporal insights, London officials gain robust evidence to regulate STLs without stifling tourism. This mirrors international efforts: New York’s Local Law 18 and Barcelona’s clampdown

on unlicensed rentals highlight the importance of data-driven policymaking. Ultimately, leveraging InsideAirbnb can lead to balanced regulations that maintain vibrant visitor economies while safeguarding long-term housing and community well-being.

References

- Barron, K., Kung, E. and Proserpio, D. (2018) 'The sharing economy and housing affordability: Evidence from airbnb', in *Proceedings of the 2018 ACM conference on economics and computation*. New York, NY, USA: Association for Computing Machinery (EC '18), p. 5. Available at: <https://doi.org/10.1145/3219166.3219180>.
- Gurran, N. and Phibbs, P. (2017) 'When tourists move in: How should urban planners respond to airbnb?', *Journal of the American Planning Association*, 83(1), pp. 80–92. Available at: <https://doi.org/10.1080/01944363.2016.1249011>.
- 'Inside airbnb' (n.d.). Available at: <http://insideairbnb.com>.
- London, airbnb market statistics & data, the united kingdom* (no date). Available at: <https://airbtics.com/annual-airbnb-revenue-in-greater-london-uk/> (Accessed: 17 December 2024).
- Meris, R. (2023) *Airbnb price optimization with airbnb historical data. Airbtics / airbnb analytics*. Available at: <https://airbtics.com/airbnb-price-optimization/> (Accessed: 16 December 2024).
- Nieuwland, S. and Melik, R. van (2020) 'Regulating airbnb: How cities deal with perceived negative externalities of short-term rentals', *Current Issues in Tourism*, 23(7), pp. 811–825. Available at: <https://doi.org/10.1080/13683500.2018.1504899>.
- Prentice, C. and Pawlicz, A. (2024) 'Addressing data quality in airbnb research', *International Journal of Contemporary Hospitality Management*, 36(3), pp. 812–832. Available at: <https://doi.org/10.1108/IJCHM-10-2022-1207>.
- Spier, S. (2024) 'Uncovering digital platforms' ethics and politics: The case of airbnb', *Philosophy & Technology*, 37(2), p. 54. Available at: <https://doi.org/10.1007/s13347-024-00742-y>.
- The insider's guide to airbnb occupancy rates in london from smarthost* (no date). Available at: <https://www.smarthost.co.uk/blog/the-insiders-guide-to-airbnb-occupancy-rates-in-london> (Accessed: 17 December 2024).
- Wachsmuth, D. and Weisler, A. (2018) 'Airbnb and the rent gap: Gentrification through the sharing economy', *Environment and Planning A: Economy and Space*, 50(6), pp. 1147–1170. Available at: <https://doi.org/10.1177/0308518X18778038>.