NOVA IMS Information Management School

MDSAA

Master Degree Program in

Data Science and Advanced Analytics

UNRAVELING THE SPATIAL DYNAMICS OF AIRBNB IN LISBON

A Data-Driven Exploration of Determinants and Insights for Urban Planning

Amanda Tavares de França

Project Work

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School

Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNRAVELING THE SPATIAL DYNAMICS OF AIRBNB IN LISBON

A Data-Driven Exploration of Determinants and Insights for Urban Planning

by

Amanda Tavares de França

Project Work presented as a partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Business Analytics.

Supervised by

Supervisor: Miguel de Castro Neto, PhD, NOVA Information Management School

Co-supervisor: Bruno Jardim, PhD, NOVA Information Management School

July, 2024

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, July 6th, 2024

ACKNOWLEDGEMENTS

I am deeply grateful to my parents, Marta and Cesar, for believing in my potential and their lifelong support. Their unconditional love has been invaluable to me. Heartfelt thanks go to my closest friends whose companionship made this journey not only bearable but enjoyable. I am especially grateful to my boyfriend, Tiago, whose admiration and motivation added meaning to every stepping stone along the way. Lastly, I extend sincere thanks to all the professors for sharing their time and expertise throughout this endeavor. To all those who have contributed to my journey, I offer my sincere appreciation.

ABSTRACT

This study adopts a data-driven approach to comprehend the determinants and factors affecting the spatial distribution of Airbnb listings across Lisbon. The primary objective is to unravel the spatial dynamics and determinants influencing Airbnb penetration, supporting urban planners and policymakers in devising more effective spatial strategies. Using spatial autocorrelation measures, the study reveals a moderate to strong positive spatial autocorrelation of Airbnb listings, with significant clustering in central neighborhoods. The Local Moran's I analysis identifies high-high clusters in central, tourist-rich areas such as Arroios and Santa Maria Maior, and low-low clusters in peripheral, less attractive neighborhoods like Benfica and Carnide.

Regression analysis highlights the key factors influencing Airbnb distribution, with proximity to the city center, density of points of interest, and accessibility to public transport emerging as significant determinants. Socio-demographic factors, particularly the proportion of young adults aged 20-39, also show a positive correlation with Airbnb listings.

The Geographically Weighted Regression model further reveals spatial heterogeneity in the relationships between Airbnb listings and independent variables. Rent prices, percentage of young adults, and number of points of interest exhibit varying impacts across different neighborhoods, underscoring the need for localized policy responses.

KEYWORDS

Spatial autocorrelation; Airbnb; Lisbon; Geographically Weighted Regression; Urban planning

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AIC	Akaike's Information Criteria
AICc	corrected Akaike Information Criterion
BIC	Bayesian Information Criteria
CDMs	Count Data Models
ESDA	Exploratory Spatial Data Analysis
GIS	Geographic Information Systems
GWR	Geographically Weighted Regression
HH	High-High Cluster
HL	High-Low Cluster
INE	National Institute of Statistics in Portugal
LISA	Local Indicator of Spatial Association
LH	Low-High Cluster
LL	Low-Low Cluster
NBM	Neighborhood-Based Model
NUTS	Nomenclature of Territorial Units for Statistics
OSM	OpenStreetMap
POI	Points of Interest
VIF	Variance Inflation Factors

1. INTRODUCTION

The ascendance of sharing economy platforms has introduced a new socioeconomic model that has transformed the production, distribution, and consumption of goods and services by allowing people to easily share underutilized resources on a fee-based sharing basis (Cesarani & Nechita, 2017). According to Fernandes (2019), the prominence of the sharing economy has grown significantly in the hospitality industry, offering opportunities for individuals and corporations to establish rapidly growing businesses.

Airbnb is a notable example of the sharing economy in the field of tourism and hospitality (Zhang & Fu, 2022). Founded in 2008 as a platform that enables short-term rentals of accommodations to tourists, Airbnb has been experiencing exponential growth and has become a dominant player in the market. According to Zhang and Fu (2022), the expansion of the platform yields numerous advantages for both users and cities. It provides additional income to hosts, boosts the tourism economies, and has a positive cultural exchange. However, there are also challenges associated with the rapid emergence of Airbnb. Its critics contend that Airbnb harms neighborhoods, undermines labor unions, and exacerbates the housing crisis (Lee & Chang, 2019).

Antunes and Ferreira (2021) state that the concept of the sharing economy conceals a truth where short-term rentals operate not only as complementary income for property owners but as unofficial hotels throughout the entire year. In this sense, platform hosts withdraw housing units from the market to use as Airbnb listings, reducing the number of available housing and increasing property prices. The unrestrained expansion of the platform listings in residential areas has also been associated with gentrification, disruption of the housing market, and increased spatial inequalities, presenting urban challenges and multiple economic and social repercussions (Singh & Azevedo, 2021).

Recently, the prominence of Airbnb has surged in discussions within political, academic, and civil society circles due to the legal and regulatory questions it poses regarding liability, taxes, zoning, and other external factors that impact urban planning (Antunes & Ferreira, 2021). Zhang and Fu (2022) urge that to efficiently regulate and monitor the expansion of Airbnb, it is essential to investigate its spatial dynamics and understand its underlying mechanisms.

Over the last decade, Lisbon, as a global tourist destination, has been dealing with various urban planning challenges due to the surge in the tourism sector, particularly the exponential growth of the short-term rental industry.

While prior studies have examined various aspects of Airbnb, such as pricing (Nunes, 2023), customer satisfaction (Pires et al., 2020), its impact on spatial inequalities (Garha, 2021), and its effect on the hotel industry (Eugenio-Martin et al., 2019), a significant gap remains regarding the factors influencing the spatial distribution of listings in Lisbon. This study adopts a data-driven approach to comprehend the determinants and factors affecting the spatial distribution of Airbnb across Lisbon. The goal is to unravel the spatial dynamics and determinants influencing the distribution of Airbnb listings, aiding urban planners and policymakers in devising more effective spatial strategies related to the platform. To achieve this, the study employs spatial autocorrelation measures, including Moran's I and Local Moran's I, as well as regression analyses such as Negative Binomial Regression and Geographically Weighted Regression (GWR). These methods help answer key questions regarding the clustering of Airbnb listings, the impact of geographic and socio-demographic factors, and the spatial variability in these relationships.

This study aims to address the following research questions:

(1) How do socio-demographic, geographic, and economic factors impact the spatial distribution of Airbnb listings in Lisbon?
(2) What are the spatial autocorrelation patterns of Airbnb distribution in Lisbon?
(3) Where are the hotspot areas of Airbnb listings in Lisbon, and how do they cluster spatially?

The thesis is organized into several sections. The Introduction provides an overview of the study's background, objectives, and significance. The Literature Review covers relevant topics, including urban planning and short-term rentals, implications of Airbnb listings' spatial dynamics, data-driven urban planning, and previous research on the spatial distribution of Airbnb listings. The Data & Methodology section describes the data sources, data processing techniques, and the methodologies used for analysis. The Results and Discussion section presents the findings of the spatial autocorrelation and regression analyses, followed by a discussion of their implications. The Conclusion and Future Works section summarizes the main findings, discusses the limitations of the study, and suggests directions for future research.

2. LITERATURE REVIEW

The literature review section delves into the existing research pertinent to the factors influencing the distribution of Airbnb listings in urban environments. Through the synthesis of recent studies and theoretical frameworks, this section seeks to identify avenues for further research.

2.1. URBAN PLANNING AND SHORT-TERM RENTALS

According to Huxley & Inch (2020), urban planning encompasses designing and managing urban areas with a focus on physical, social, and economic factors. Its objective is to promote sustainable cities that are both livable and functional by addressing issues such as land usage, transportation, housing availability, public spaces, and environmental sustainability. It entails formulating plans and policies that guide land use, providing infrastructure and services, and constructing buildings alongside public spaces. Urban planning involves collaboration among citizens, government officials, urban planners, and developers to incorporate the perspectives of all stakeholders (Medeiros, 2021).

Tourism can have a significant impact on urban planning. Borges et al. (2022) state that while it can enhance the quality of life for residents and attract visitors through improved amenities, it also has the potential to trigger gentrification, displacement of locals, and cultural heritage erosion.

Estevens et al. (2023) described that during the early 2000s in Lisbon, there was a shift in the local state's goals for the city, emphasizing positioning it as a global city. This resulted in policies that favored private developers over urban planning considerations. The 2012 Municipal Master Plan categorized a significant portion of the city as an "urban rehabilitation area" and provided fiscal incentives to neighborhoods beyond the central districts. This strategy aimed to attract international investment into the local real estate market and led to a surge in housing rehabilitation projects, mainly catering to short-term rentals and upscale residences for international users.

This trend of urban rehabilitation prioritizing private capital interests over the needs of the local population resulted in a rapid increase in short-term listings in Lisbon and raised concerns about its effects on urban planning (Estevens et al., 2023).

The challenges confronting Lisbon's metropolitan area in terms of urban planning due to the surge in short-term rentals are multifaceted. Notably, concerns revolve around the adverse effects on housing availability and affordability, as well as the increased demand for transportation services driven by tourism (Louro et al., 2021).

At the heart of the short-term rental surge lies Airbnb, a global platform founded in 2008 that facilitates the renting of entire homes or individual rooms to travelers. Operating in over 220 countries and regions, the platform has emerged as a major player in the sharing economy, characterized by digital technologies facilitating peer-to-peer transactions (Petruzzi et al., 2020). Estevens et al. (2023) outline that its role as a central element in the touristification of places is underscored by its ability to promptly provide millions of accommodation options in areas initially not intended for tourism.

2.2. Implications of Airbnb listings' spatial dynamics

Singh and Azevedo (2021) state that the expansion of Airbnb has disrupted the conventional tourism industry by introducing an alternative form of accommodation that is frequently more affordable and flexible than traditional hotels. This shift brings about both favorable and adverse effects on local tourism dynamics, with Airbnb's spatial distribution having the potential to reshape tourism dynamics by altering the array of accommodations available to visitors and, potentially, increasing the overall number of tourists (Zhang & Fu, 2022).

The consequences of this touristification in Lisbon are evident in the form of gentrification, leading to the displacement of residents and the transformation of neighborhoods into tourist-centric enclaves (Estevens et al., 2023). As stated by Sun et al. (2021), the spatial distribution of short-term rentals plays a pivotal role in altering neighborhood dynamics, influencing the composition of residents and visitors within a specific area, and raising conflicts between them. This shift can lead to a redefinition of the social fabric, with implications for community interactions and cohesion. This phenomenon becomes particularly pronounced when tenants and homeowners are displaced (Estevens et al., 2023).

Singh and Azevedo (2021) state that the concentration of Airbnb listings in specific neighborhoods exacerbates spatial inequalities and overcrowding in major tourist destinations. According to Quattrone et al. (2018), this spatial distribution has far-reaching consequences, influencing urban space utilization, resource allocation, tourism infrastructure development, and public service delivery. It imposes significant challenges on city management, compelling them to address issues ranging from noise complaints to waste management and safety concerns.

The surge of Airbnb listings has also been associated with a decline in the availability of long-term rental housing, resulting in gentrification, and elevated rents and housing prices (Singh & Azevedo, 2021). Research conducted by Franco and Santos (2021) underscores the significant influence of Airbnb on the values of residential properties and rental rates in Portugal. It reveals that a 1 percentage point increase in a municipality's Airbnb share is associated with a 0.018% increase in house prices and a 0.028% increase in rents. This empirical evidence emphasizes the relationship between Airbnb, urban planning, and the changing socio-economic landscape of Lisbon. Consequently, discussions have emerged on the necessity to regulate and monitor short-term rentals in urban areas, emphasizing the need to strike a balance between the interests of tourists and residents (Quattrone et al., 2018). As maintained by Ki and Lee (2019), employing spatial analysis becomes crucial to comprehend the impact of Airbnb on housing markets and its influence on the availability and affordability of housing in specific areas.

The Airbnb phenomenon is inherently spatial, and understanding it is paramount for policymakers seeking to harmonize its growth with broader urban planning objectives and allows for the formulation of targeted policies. Such policies can seek to balance the interests of hosts, tourists, and local communities, mitigating potential negative impacts on housing markets while fostering sustainable urban development (Estevens et al., 2023).

2.3. DATA-DRIVEN URBAN PLANNING

The field of urban planning has entered a new phase with the introduction of data-driven methods that utilize advanced analytical techniques. For instance, urban planning benefits from Geographic Information Systems (GIS), as highlighted by Medeiros et al. (2021) and Santos et al, 2022. GIS can be used to manage and model spatial connections among various urban attributes, including population density, land use, street connectivity, commercial diversity, and urban structure (Santos et al., 2022).

Santos et al. (2022) explain that the integration of big data and machine learning algorithms further empowers urban planners to analyze and predict complex phenomena such as traffic congestion, air pollution, and energy consumption, facilitating the design of targeted interventions. The study by Lee and Chang (2019) mentions that the application of forest-based classifications offers a robust approach to identifying key drivers of urban phenomena. This technique, applied to large datasets, unveils patterns influencing the spatial distribution of elements such as short-term rentals. By deciphering these patterns, planners can enhance their understanding of factors influencing urban development, paving the way for more effective policy design (Estevens et al., 2023).

These approaches not only reveal complex urban dynamics but also lead to forward-thinking, adaptable, and sustainable urban development. Incorporating data-driven analysis with conventional planning techniques signifies a significant change, guaranteeing that future urban planning is both well-informed and flexible to the changing requirements of diverse and dynamic urban societies.

2.4. RELEVANT RESEARCH ON SPATIAL DISTRIBUTION OF AIRBNB LISTINGS

The exploration of spatial distribution in the context of Airbnb listings has been studied significantly through advanced methods that combine spatial analysis and machine learning techniques, summarized in table 2.4.

Table 2.4 – Summary of Studies on Spatial Distribution of Airbnb Listings

Study	Objective	Data	Method	Findings
Lee and Chang (2019)	Identify high-demand areas for Airbnb listings in Los Angeles and understand factors influencing tourist spatial behavior.	Airbnb open-source data (2016). Information on Airbnb listings, consumer behavior, and host characteristics in Los Angeles.	Geospatial analysis, Random Forest Classification, Forest-Based Classification and Regression	Tourist spatial behavior is influenced by location, price, amenities, and group size. Economic effects include increased housing costs, gentrification, and potential elevation of consumer prices.
Eugenio-Martin et al. (2019)	Analyze spatial distribution of Airbnb listings in the Canary Islands, focusing on different tourist destinations.	Instituto Canario de Estadística (ISTAC) data (2018) for hotels and apartments. Airdna dataset for Airbnb properties (May 2017). Tourist visit data from 194,777 unique tourism locations (2005 to March 2018).	Generalized two-stage least squares, bivariate spatial correlation	Airbnb distribution determinants vary by destination type. Pricing, population, and tourist visits identified as significant factors.
Singh and Azevedo (2021)	Investigate the impact of Airbnb on urban built environment in selected Spanish cities, exploring spatial relationships.	Inside Airbnb data for Barcelona and Lisbon (Sep 2019 and Dec 2020). 2011 Census for socioeconomic and demographic data. Municipal registers (2011-2019) for population and housing information. Open Street Maps data for transport access.	Regression models, GIS, Local Indicator of Spatial Association index	Airbnb penetrates central residential zones and major tourist destinations, exacerbating spatial inequalities, particularly favoring higher-income homeowners.
Antunes and Ferreira (2021)	Scrutinize spatial distribution of short-term rentals in Barcelona and Lisbon, identifying determinants and impacts.	*Registo Nacional de Turismo* (RNT) and *Sistema de Informação Geográfica do Turismo* (SIGTUR) data. Spatial distribution data for Lisbon municipality.	Spatial and multiscale analysis, GIS, standard distance, standard deviational ellipse, clustering hotspots	Determinants include distance to city center, tourist points, population density, housing sizes, household income, and immigrant share.

Table 2.4 – Summary of Studies on Spatial Distribution of Airbnb Listings (Continued)

Study	Objective	Data	Method	Findings
Sun et al. (2022)	Analyze factors influencing spatial distribution of Airbnb listings in Suzhou, China, using a spatial econometric model.	POI data from AirDNA (Nov 2017-2020). GIS data for Suzhou. On-site interviews with residents, tourists, and B&B operators.	Spatial bivariate correlation, geography-weighted regression analyses	Spatial distribution is influenced by factors like urban location, catering services, proximity to tourist attractions. High concentration in the old city and near the railway station.
Zhang and Fu (2022)	Examine Airbnb supply in Los Angeles, considering neighborhood characteristics, proximity to attractions, and socioeconomic factors.	Airbnb data for Los Angeles (2014 and 2019). Socioeconomic data on resident population, housing units, income, and rent. Spatial datasets for Los Angeles.	Exploratory spatial data analysis (ESDA), spatial autocorrelation analysis, spatial regression models	Factors affecting spatial distribution include resident population, housing units, income, rent, points of interest, and proximity to Hollywood.

These methods provide valuable insights into the various factors impacting and the resulting consequences of the observed spatial patterns.

Lee and Chang (2019) utilize geospatial analysis and machine learning algorithms, including Random Forest Classification and Forest-Based Classification and Regression, to pinpoint high-demand areas for Airbnb listings in Los Angeles. Overall, the study's findings suggest that tourist spatial behavior in Los Angeles is influenced by a combination of factors, including location, price, amenities, and group size. Lee and Chang emphasize several significant discoveries, including the debated economic effects of Airbnb. The platform has been linked to increased housing costs, the gentrification of neighborhoods, and the potential elevation of consumer prices for goods and services.

Eugenio-Martin et al. (2019) contributes a spatial econometric model to analyze the spatial distribution of Airbnb listings in the Canary Islands. Through the application of generalized two-stage least squares and bivariate spatial correlation, the study aims to comprehend the diverse spatial connections between traditional hotels and Airbnb in various categories of local tourist destinations, specifically those focused on sun and beach, nature-based, and city tourism. Their analysis reveals that the determinants of Airbnb distribution vary depending on the type of destination. The most significant factor identified for Airbnb entry is pricing, followed by population and tourist visits, which play significant but slightly less prominent roles. The study also emphasizes the importance of considering the impact of Airbnb spatial distribution on policy development and tourism dynamics, particularly in the context of different types of tourism destinations.

Singh and Azevedo (2021) focused on selected cities in Spain to employ a combination of regression models and GIS. Using the Local Indicator of Spatial Association (LISA) index and regression analysis, they investigate the impact of Airbnb on urban built environment elements, providing insights into the spatial relationship between Airbnb housing sources and various urban factors. In particular, the research revealed that Airbnb has penetrated deeply into central residential zones and surrounding major tourist destinations, exacerbating spatial inequalities by mainly favoring homeowners from the higher-income group in these neighborhoods.

Antunes and Ferreira (2021) take a different approach, leveraging spatial and multiscale analysis to scrutinize the spatial distribution of short-term rentals in Barcelona and Lisbon. Using GIS and spatial statistics tools, their study incorporates techniques such as standard distance, standard deviational ellipse, and clustering hotspots to reveal spatial patterns and trends. The paper identifies several determinants of the spatial distribution of Airbnb properties, including the distance to the city center, concentration of POI for tourists, low population density, housing sizes, household income, and share of immigrants. Additionally, the study highlighted the notable concentration of short-term rental units in urban regions, specifically in Lisbon. It discussed the positive and negative impacts these concentrations exert on urban life and underscored potential concerns such as gentrification, housing financialization, and the transformation of places into tourist enclaves.

Sun et al. (2022) introduced a spatial econometric model to analyze the factors influencing the spatial distribution of Airbnb listings in Suzhou, China. Using various statistical techniques, including spatial bivariate correlation and geography-weighted regression analyses, the study investigated the relationships between Airbnb's spatial distribution and factors such as tourist attractions, transportation accessibility, and the neighborhood environment. The findings suggested that the spatial distribution of Airbnb in Suzhou is influenced by multiple factors, including the urban location, the availability of catering services, and proximity to tourist attractions. The study revealed a high concentration of listings within the old city and near the railway station, indicating the significance of these areas. Additionally, the attributes of Airbnb listings in Suzhou were discovered to be closely associated with city centers, historical neighborhoods, and traditional residences. This dependence on specific locations raises concerns, particularly in the historic district, where the influence of Airbnb on historic preservation and community dynamics is a significant issue.

Zhang and Fu (2022) present a comprehensive methodology in their examination of Airbnb supply in Los Angeles employing a combination of exploratory spatial data analysis (ESDA), spatial autocorrelation analysis, and spatial regression models. Their study encompasses factors such as neighborhood characteristics, proximity to tourist attractions, and socioeconomic factors and identified several factors affecting the spatial distribution of Airbnb supply, such as resident population, housing units, income, rent, points of interest (POI), and proximity to Hollywood. According to the authors, these findings have implications for urban planning, city management, and regulation of short-term rentals, emphasizing the need for efficient and effective management strategies to address the increasing and geographical clustering trend of Airbnb supply in tourism cities, as well as the potential impact on housing markets, neighborhood dynamics, and tourism dynamics.

These diverse approaches, from spatial analysis to machine learning and econometric models, offer nuanced insights into the intricate interplay of factors influencing the spatial dynamics of Airbnb rentals. This geographic approach is pivotal for analyzing their distribution across cities like Lisbon, enabling the implementation of tailored urban planning measures at various scales, from districts and regions to specific neighborhoods or streets.

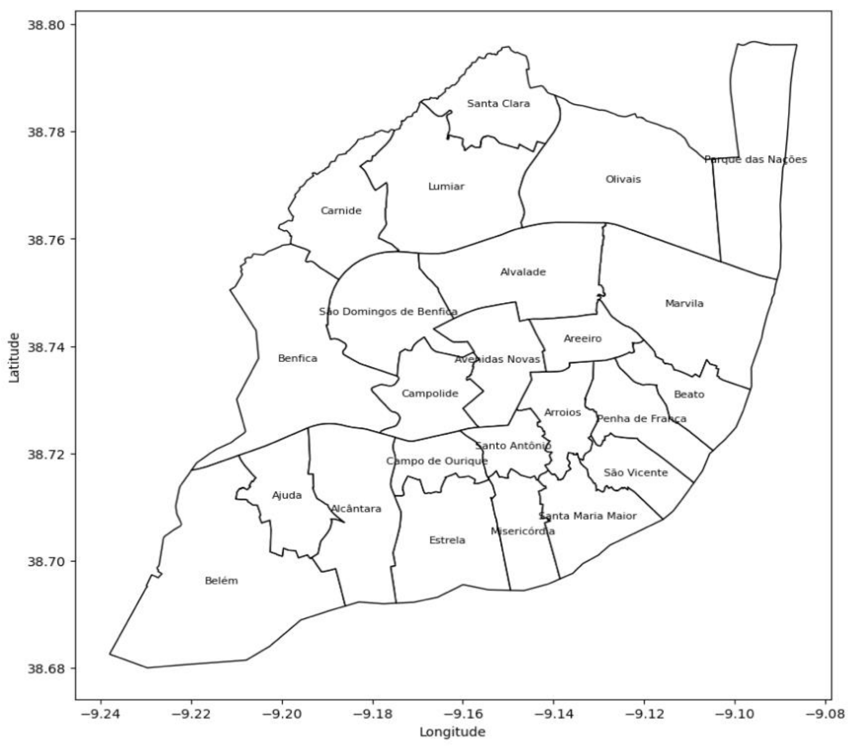
3. DATA & METHODOLOGY

In this section, there's a comprehensive overview of the data utilized and the methodologies employed in this study. It begins by describing the data sources and the key variables involved, followed by an outline of the preprocessing steps undertaken to prepare the data for analysis. Subsequently, it details the statistical techniques used to analyze the data, emphasizing the rationale behind the choice of each method. This section aims to ensure a clear understanding of the data foundation and the analytical framework that support the findings and conclusions.

3.1. DATA

The focal point of this study is the city of Lisbon, acclaimed as the westernmost capital of the European continent, nestled along the Atlantic coast. Serving as a pivotal economic hub within Europe, Lisbon stands as one of its principal centers of commerce and industry. It is among the most sought-after destinations in Europe, having already won the awards for "Best City Destination" and "Best City Break Destination" at the esteemed World Travel Award (Área Metropolitana de Lisboa, 2023). Spanning across an expansive area of 100 km2, Lisbon is home to over 500 thousand residents, constituting approximately 3% of the country's total population (Lisbon City Council, 2023). The city encompasses 24 parishes, as depicted in the map presented in figure 3.1.1.

Figure 3.1.1 – Map of the City of Lisbon Divided by Neighborhoods



Source: Lisbon City Council/Georeferenced Information Management Division (2023)

To better understand the factors influencing the spatial distribution of Airbnb in Lisbon, different sets of variables were used divided into Airbnb, Geographic, Socio-economic and Demographic data, described in table 3.1.1.

Table 3.1.1 – Exploratory Variables Used in the Analysis

Data Category and Variable	Description
Airbnb	
Listings	Number of listings in the neighborhood.
Geographic	
Distance to Center	Distance between the neighborhood centroid and city center.
Transportation	Number of bus, metro and train stops by geographic location.
POI	Number of points of interest by geographic location.
Socio-Economic	
Unemployment Rate	Percentage of unemployed individuals by Place of Residence.
Rent Value	Average value of rents per square meter of new lease contracts (2023) for family accommodations.
Education Level	Proportion of resident population with completed higher education (%) by Place of Residence.
Demographic	
Population Density	Population density (N.º/km²) by Place of Residence.
Young Adults	Number of residents between the ages of 20-39 by Place of Residence.
Housing Units	Number of rented classic family accommodations for habitual residence by geographic location.

Airbnb data was retrieved from Inside Airbnb, a project with a mission to provide data and advocacy regarding the impact of Airbnb on residential communities. Founded by Murray Cox, an artist, activist, and technologist, the project aims to empower communities with information to help them understand, make decisions, and control the role of renting residential homes to tourists. The website utilizes various Open-Source technologies, including D3, Bootstrap, Python, PostgreSQL, and Google Fonts (Inside Airbnb, 2023).

The data obtained from Inside Airbnb includes details on host and room identification through a unique key, accommodation location determined by the neighborhood, room type (entire homes, private rooms, or shared rooms), geographical coordinates (latitude and longitude), accommodation specifics such as the number of bedrooms and bathrooms, maximum guest capacity, and minimum stay period, pricing, and user reviews available on the website. For this project, data from Inside Airbnb was collected from January to December 2023 for the city of Lisbon.

The spatial data was retrieved from OpenStreetMap (OSM), a collaborative and open-source mapping platform launched in 2004. Volunteers, contributors, and organizations use GPS devices, aerial imagery, and field surveys to collect and update information about roads, trails, buildings, and various geographic features (OpenStreetMap, 2024).

The data retrieved from OSM encompasses a wide array of attributes pertinent to geographical elements such as amenities, transportation infrastructure, buildings, and historical landmarks. This includes information on shops, tourism sites, and specific amenities like supermarkets, restaurants, fast food outlets, cafes, bars, and pubs. Moreover, the data encompasses polygons characterized by leisure parks as well as significant historical sites. Additionally, the dataset comprises polygons featuring aeroway, highway, railway, or specific amenities such as ferry terminals or bus stations, along with buildings such as aerodromes, ferry terminals, or train stations.

The demographic, social, and economic variables were obtained from the National Institute of Statistics in Portugal (INE), the country's official statistical agency. Operating under the Ministry of Finance, INE is tasked with collecting, analyzing, and disseminating comprehensive and accurate statistical information across various domains, including demographics, economics, and social indicators. Through surveys, censuses, and collaboration with international organizations like Eurostat, INE ensures the production of reliable data that supports evidence-based decision-making, policy formulation, and academic research (INE, 2024).

The data retrieved from INE provides information on various socio-economic and demographic indicators from the 2021 Census, categorized by geographic location (NUTS - 2013). Data encompasses the number of housing units, average value of rents, population density, level of education, age structure, and unemployment rate by parish.

3.2. METHODOLOGY

In accordance with Tobler's (1970) first law of geography, which states that "Everything is related to everything else, but near things are more related than distant things", spatial autocorrelation analysis was utilized to grasp the tendency of adjacent areas to exhibit similar attributes or values.

Spatial autocorrelation refers to the inclination of neighboring areas to share similar characteristics or values, revealing a phenomenon where a variable in one location correlates with the same variable in nearby locations within a two-dimensional space (Kim & Lee, 2021). Lee et al. (2012) suggested that landscape features demonstrate spatial autocorrelation due to the continuity and hierarchical structure of environmental conditions. This concept is fundamental in geography and spatial analysis, indicating that spatial patterns exhibit organization or clustering rather than random distribution (Kim & Lee, 2021).

To assess spatial autocorrelation, the Moran's I test statistic, introduced by Cliff and Ord (1981), was employed. This statistic provides insights into overall autocorrelation and helps in understanding the spatial distribution and structure of autocorrelation patterns. It serves as a tool for inference, operating under the hypothesis that variable values are randomly distributed across space. The test yields a statistic ranging from -1 to +1 where higher positive values indicate spatially adjacent areas with similar data values, 0 indicates randomness, and closer to -1 signifies a mix of high and low data values (Kim et al., 2022).

The Global Moran’s I was computed as follows:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \, (y_i - \bar{y})(y_j - \bar{y})}{s_0 \sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{1}$$

where $(y - \underline{y})$ is the deviation of an attribute for a feature from its mean, w_{ij} is the spatial weight between features and n is the total number of features, and s_0 is the sum of all spatial weights (Sarkar et al., 2024).

While the global Moran’s I index provides insights into the overall spatial autocorrelation across areas under study, it lacks the ability to discern localized patterns of spatial association (Kim & Lee, 2021). To address this limitation, Anselin (1995) introduced the LISA index, designed specifically for measuring spatial association at the local level. LISA or Local Moran’s I, determines spatial autocorrelation by comparing the weighted average between the values of a particular area and those of its adjacent areas.

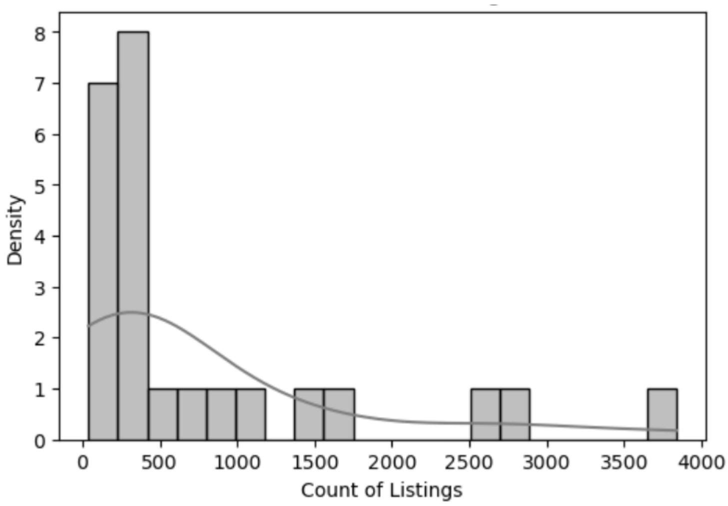
The local value of a LISA is computed as follows:

$$I_{local} = \frac{(y_i - \bar{y})}{s^2} \sum \mathrm{wij(yi - \ y)} \tag{2}$$

where w_{ij} represents the weight matrix indicating proximity relationships between different spatial units, y represents the attribute values of spatial units, $\underline{y}$ is the average number of listings, and s^2 is the variance in the number of listings in different territorial units. A local Moran’s I greater than 0 signifies spatial clustering, where high values are surrounded by high values (HH) or low values are surrounded by low values (LL). Alternatively, a value smaller than 0 indicates spatial outliers, where low values are surrounded by high values (LH) or high values are surrounded by low values (HL) (Singh & Azevedo, 2021).

To uncover the main factors affecting the spatial penetration of Airbnb listings, regression analysis was employed. Given that the dependent variable, the number of listings, is a non-negative integer and follows a Poisson distribution as shown in Figure 3.2.1, Count Data Models (CDMs) were applied.

Figure 3.2.1 – Airbnb Listings Distribution



The local spatial correlation analysis revealed significant spatial clustering of listings in specific areas of Lisbon. To address the issue of overdispersion, a Negative Binomial Regression Model (NBM), popularized and extensively discussed by Hilbe (2011), was employed.

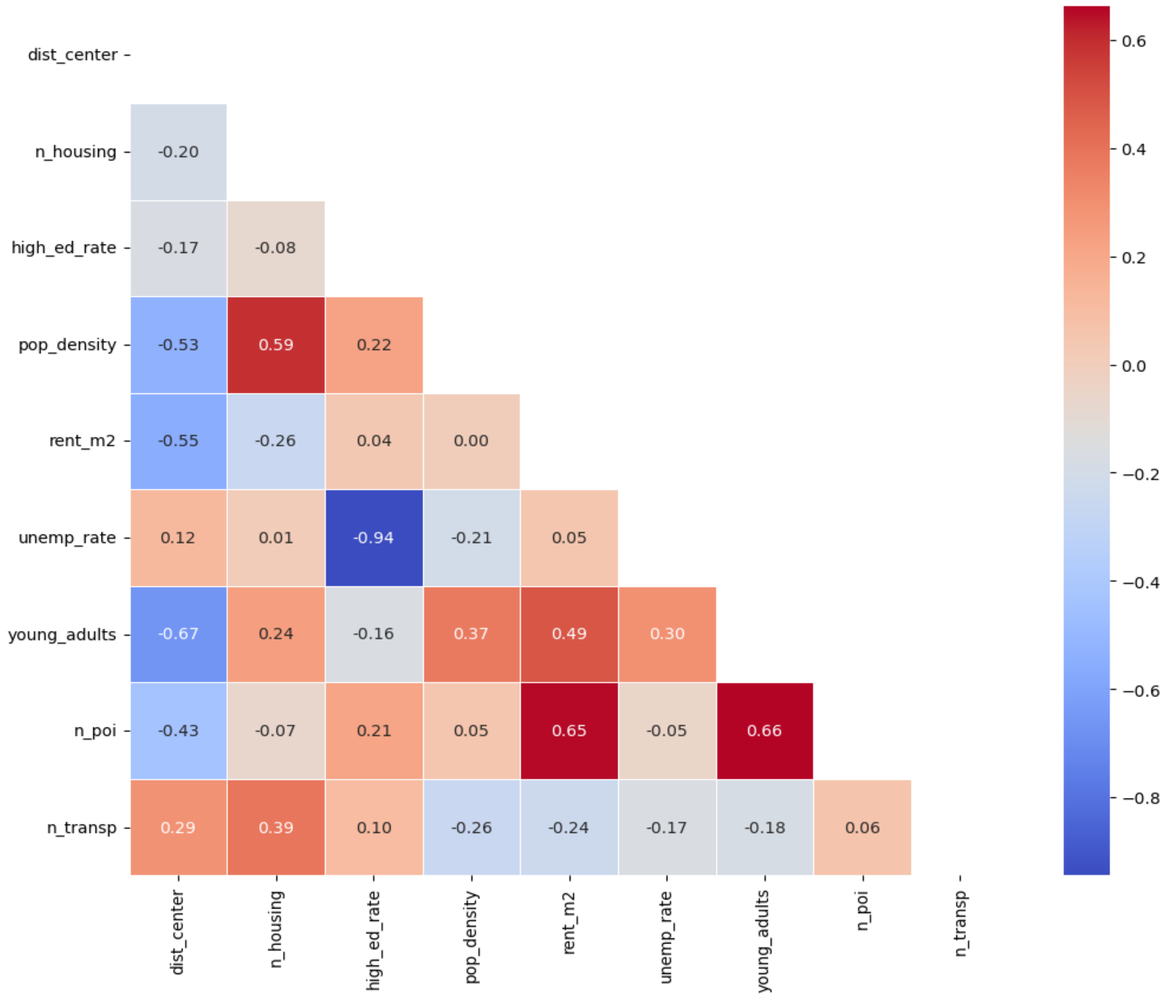
The NBM regression equation is represented as follows:

$$log\big(E(y_i)\big) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} \tag{3}$$

where $E(y_i)$ is the expected count of the dependent variable, β_0 is the intercept term, $\beta_1, \beta_2 \ldots, \beta_k$ are the coefficients of the independent variables, and $X_{i1}, X_{i2}, \ldots, X_{ik}$ are the independent variables.

Additionally, the correlation between explanatory variables was examined, and multicollinearity was controlled by ensuring low levels of Variance Inflation Factors (VIF). The correlation of independent variables can be seen in Figure 3.2.2. To address the issue, two negative binomial regression models were implemented. The first model included all variables, while the second model excluded highly correlated variables. This approach ensures that the effect of each independent variable on the dependent variable is accurately captured and allows for the selection of the model with the best goodness-of-fit indicators which was supported by results from Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC).

Figure 3.2.2 – Independent Variables Pearson Correlation Heatmap



Geographically Weighted Regression (GWR), introduced by Brunsdon, Fotheringham, and Charlton (1996), was applied to identify variations in predictor strength across different locations. Unlike NBM, which fits a single regression equation across the entire study area, GWR builds a local equation for each feature in the dataset. It constructs these separate equations by incorporating the dependent and explanatory variables of the features within the neighborhood of each target feature.

A global regression model is created first by fitting a single regression equation using all data points across the entire study area, ignoring spatial heterogeneity. This global model assumes that relationships between variables are constant across space. In contrast, the local regression models in GWR are created by fitting separate regression equations for each location within the study area, incorporating spatially varying relationships by weighting nearby observations more heavily than distant ones. The resulting local models capture spatially specific variations in relationships between variables, allowing for a detailed comparison with the global model to assess how spatial heterogeneity influences the regression outcomes across different locations within the study area.

The chosen kernel type selected was Gaussian, which is more effective when there is a notable spatial pattern in the data since it allows for capturing gradual variations in Airbnb listings across neighborhoods while incorporating any spatial clustering tendencies. The kernel selection method used was the Golden Search, based on minimizing the value of the corrected Akaike Information Criterion (AICc). This method first determines the maximum and minimum distances and then tests the AICc at various distances incrementally between them. The maximum distance is the point at which every feature has half the number of input features as neighbors, and the minimum distance ensures every feature has at least 5 percent of the dataset's features as neighbors (Esri, 2024). The GWR model can be written as:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{K} \beta_k(u_i, v_i) X_{ik} + \epsilon_i \tag{4}$$

where y_i is the dependent variable, $\beta_0(u_i, v_i)$ is the intercept term at location i, (u_i, v_i) are the geographical coordinates, $\beta_k(u_i, v_i)$ are the coefficients for the k-th independent variable at location i, X_{ik} are the explanatory variables and ϵ_i is the error term with zero mean and homogenous variance.

4. RESULTS AND DISCUSSION

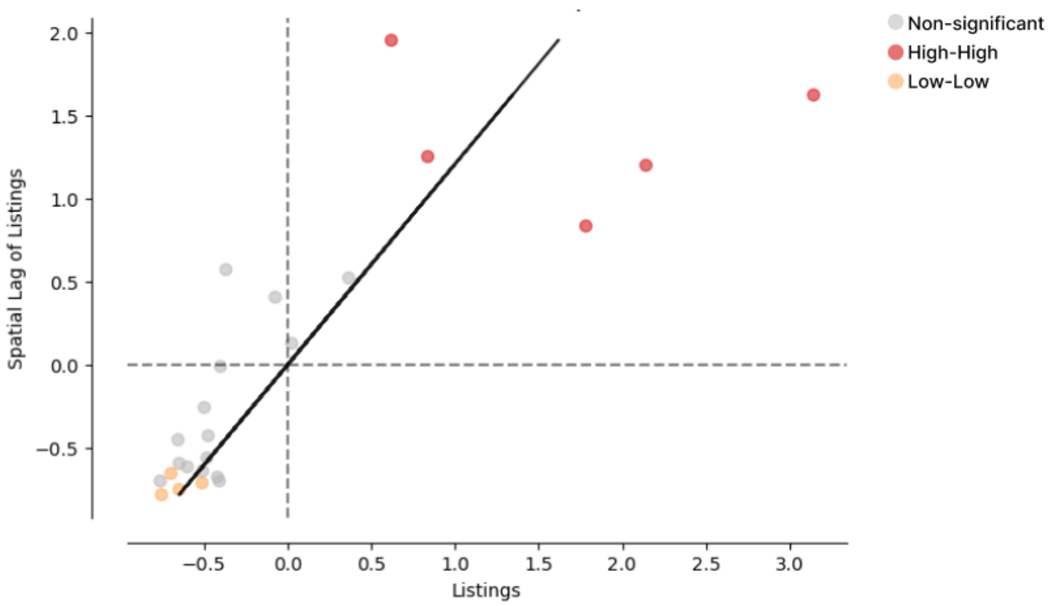
This section presents the findings of the study and discusses their implications. The analysis begins with an examination of spatial autocorrelation to understand the clustering patterns of Airbnb listings in Lisbon. Following this, the determinants of Airbnb's spatial distribution are analyzed through regression models to identify significant geographic and socio-demographic factors. Lastly, the Geographically Weighted Regression (GWR) model is employed to explore spatial variability in these relationships, highlighting how the impact of various factors differs across neighborhoods.

4.1. SPATIAL AUTOCORRELATION

The Moran's I statistic of 0.56 suggests that there is a moderate to strong positive spatial autocorrelation of Airbnb listings in Lisbon neighborhoods, and this spatial pattern is statistically significant with a p-value of 0.001. It implies that neighborhoods with similar numbers of Airbnb listings tend to be spatially clustered.

The Local Moran's I analysis unveiled two distinct spatial clusters: HH and LL, illustrated in Figure 4.1.1.

Figure 4.1.1 – Plot of Local Moran's Autocorrelation

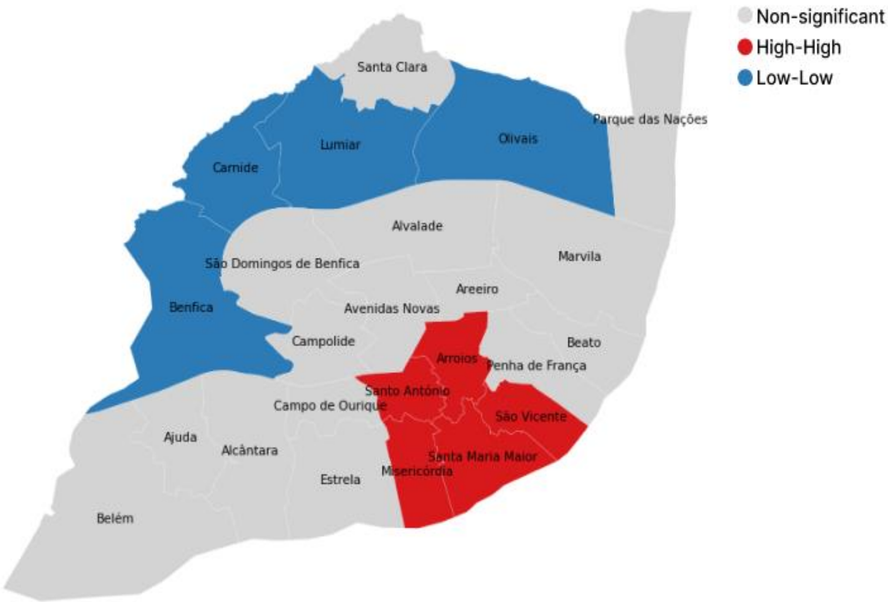


HH clusters, located in the upper right quadrant of the scatterplot, indicate neighborhoods characterized by both substantial Airbnb listing counts and pronounced spatial autocorrelation with adjacent areas. Conversely, LL clusters located in the lower left quadrant of the scatterplot, comprise neighborhoods with minimal Airbnb activity and display analogous spatial clustering tendencies.

As seen in Figure 4.1.2, in 2023 the HH cluster encompassed Arroios (13.4%), Misericórdia (15.3%), Santa Maria Maior (20.5%), Santo António (8.5%), and São Vicente (7.3%), collectively hosting 65% of Lisbon's Airbnb listings. These local units are distinguished by a high number of POI and proximity to the city center, which draw a significant inflow of visitors.

Benfica (0.5%), Carnide (0.2%), Lumiar (0.7%), Olivais (0.1%), and São Domingos de Benfica (0.1%) accounted for less than 4% of Lisbon's total listings and were classified as LL clusters. These neighborhoods are marginalized and offer little to no touristic attractions.

Figure 4.1.2 – Map of Local Moran's Autocorrelation



The identification of these spatial clusters has significant implications for urban planning and policy-making. The HH clusters, with their high concentration of Airbnb listings, highlight areas of potential strain on local infrastructure and housing markets, necessitating targeted regulatory measures to manage the impact on residents and maintain neighborhood character. Conversely, the LL clusters indicate areas that may benefit from tourism development initiatives to boost local economies and attract investment. Understanding these spatial patterns allows for more informed decisions regarding the distribution and regulation of short-term rentals in Lisbon, ensuring balanced urban development and equitable resource allocation across neighborhoods.

4.2. DETERMINANTS OF SPATIAL DISTRIBUTION

The regression analysis reveals significant insights into the geographic and socio-demographic factors influencing Airbnb penetration, as summarized in Table 4.2.1. Firstly, there is a notable negative correlation between distance from the city center and the number of Airbnb listings, indicating fewer listings as distance increases from the central area. Additionally, the presence of points of interest (POI) positively impacts Airbnb penetration, suggesting that areas with tourist attractions attract more short-term visitors. Furthermore, the accessibility of public transport services plays a crucial role in shaping the distribution of Airbnb listings.

Regarding socio-demographic factors, to mitigate multicollinearity issues identified in the initial model—specifically, variables such as the proportion of residents with higher education and the unemployment rate, which exhibited VIF values exceeding 3—one of these variables was omitted in the subsequent model. This adjustment increased the significance of rent while reducing the significance of the high education rate.

Moreover, there is a statistically significant positive association between the concentration of individuals aged 20-39 and the number of Airbnb listings in Lisbon. This finding suggests that neighborhoods with a higher proportion of this demographic tend to host more Airbnb accommodations.

Table 4.2.1 – Results of Negative Binomial Regression

Coefficient	**1st Model**	**Significance**	**VIF**	**2nd Model**	**Significance**
(Intercept)	1.4373			1.9257	*
Geographic					
Distance to Center	-0.0709	.***	1.97	-0.2844	***
Transportation	0.0073	***	1.24	0.0034	**
POI	0.0001	***	1.53	0.0009	***
Socio-Economic					
Unemployment Rate	-11.2244	*	4.77		
Rent Value	0.1952	*	1.43	0.2137	**
Education Level	0.5923	**	4.82	1.5856	.*
Demographic					
Population Density	0.0052	*	1.08	0.0087	
Young Adults	15.6744	***	1.54	12.7309	***
Housing Units	0.0572	*	1.23	0.0002	
Goodness of Fit					
AIC	324			248	
BIC	90			13	

Legend: Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The regression analysis underscores the multifaceted nature of Airbnb distribution in Lisbon. Proximity to the city center and points of interest are strong determinants, highlighting the attractiveness of central, well-serviced, and touristic areas for short-term rentals. Public transport accessibility further emphasizes the importance of connectivity in enhancing Airbnb penetration. Socio-demographically, the prominence of younger adults in driving Airbnb listings points to the appeal of such platforms for a younger, more mobile demographic. The adjustments made to account for multicollinearity also reveal the nuanced interplay between socio-economic factors, such as rent and education levels, in influencing the spatial dynamics of Airbnb listings. These findings provide a comprehensive understanding of the factors shaping Airbnb's presence in Lisbon, offering valuable insights.

4.3. GEOGRAPHIC WEIGHTED REGRESSION

The results from the GWR demonstrate nuanced spatial variability in the relationship between the number of Airbnb listings and its independent variables compared to the global regression model, as summarized in Table 4.3.1.

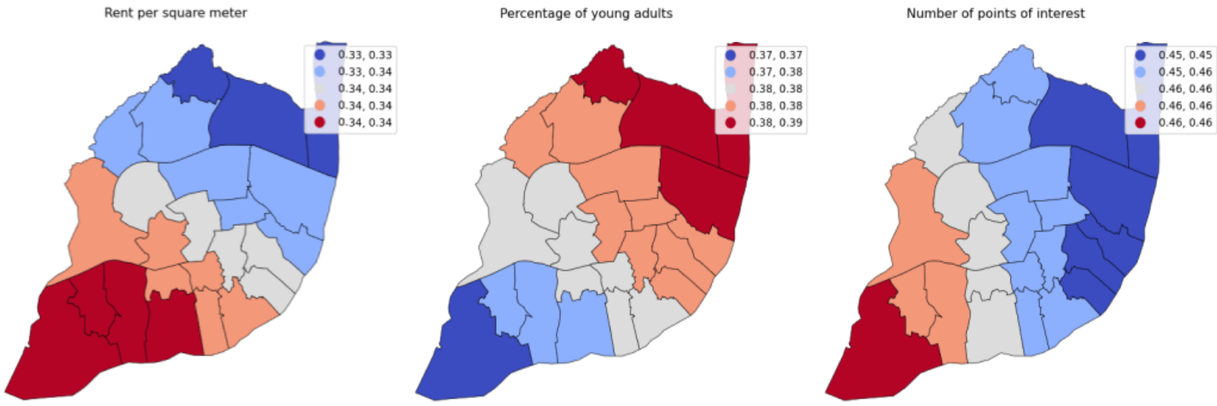
Table 4.3.1 – Statistically significant Independent Variables

Goodness of Fit Indicators	Global Regression Results	Local Regression Results
Residual Sum of Squares	1.029	1.002
AICc	36.530	39.740
BIC	-43.463	28.384
R2	0.957	0.958
Adjusted R2	0.930	0.923

The GWR model, employing a Gaussian kernel with a bandwidth of 26.870, shows marginal improvements in key diagnostic measures such as the residual sum of squares (1.002) and adjusted R-squared (0.923) compared to the global model. (Residual sum of squares: 1.029, Adj. R-squared: 0.930).

The GWR results highlight that only three independent variables are statistically significant at the local level, based on their t-values. These variables are rent by square meter, percentage of young adults, and number of POI. These findings are crucial as they indicate varying impacts of these factors across different areas within the study region, as illustrated in Figure 4.3.1.

Figure 4.3.1 – Map of Statistically Significant GWR Coefficients



Contrary to the global model, which suggested uniform relationships between variables and the number of listings across Lisbon, the GWR reveals spatial heterogeneity. This implies that the influences of rent by square meter, percentage of young adults, and number of POI on Airbnb listings differ significantly across neighborhoods or zones within the city.

5. CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH

This study examined the spatial distribution of Airbnb listings in Lisbon, focusing on the impact of socio-demographic, geographic, and economic factors. Using advanced statistical methods, including Negative Binomial regression and Geographically Weighted Regression (GWR) models, the analysis identified key determinants influencing Airbnb penetration. Additionally, spatial autocorrelation techniques, such as Local Moran's I, were employed to detect clustering patterns of Airbnb activity across different neighborhoods. The research also pinpointed hotspot areas and explored the spatial variability in the relationships between independent variables and the distribution of Airbnb listings.

The research questions are answered by this study as follows:

(1) Spatial Autocorrelation Patterns of Airbnb Distribution

The analysis confirmed a significant spatial autocorrelation among Airbnb listings across Lisbon's neighborhoods, indicating pronounced clustering tendencies. This finding is crucial for urban planners and policymakers as it highlights areas with high and low Airbnb activity, which can inform targeted interventions and regulatory strategies. The presence of spatial autocorrelation suggests that the distribution of Airbnb listings is not random but influenced by underlying geographic and socio-economic factors that create clustering effects.

(2) Identification and Clustering of Airbnb Hotspot Areas

The Local Moran's I analysis identified High-High (HH) and Low-Low (LL) clusters, illustrating distinct spatial patterns. HH clusters are predominantly located in central and touristic neighborhoods such as Arroios, Misericórdia, and Santa Maria Maior. These areas benefit from their proximity to points of interest (POI) and the city center, making them attractive for Airbnb activity. Conversely, LL clusters in peripheral neighborhoods like Benfica and São Domingos de Benfica highlight areas with minimal tourist appeal and lower Airbnb penetration. These clusters underscore the spatial heterogeneity in Airbnb listings distribution, reflecting varying levels of attractiveness and connectivity across different parts of the city.

(3) Impact of Socio-Demographic, Geographic, and Economic Factors on Airbnb Listings Distribution

The spatial distribution of Airbnb listings in Lisbon is significantly influenced by a combination of socio-demographic, geographic, and economic factors. The regression analysis, particularly through the Negative Binomial and GWR models, revealed that proximity to the city center, accessibility to public transport, and the density of POI are positively correlated with the number of Airbnb listings. These factors indicate that areas with higher tourist attractiveness and better connectivity tend to host more

Airbnb listings. Socio-demographic variables, such as the proportion of young adults (aged 20-39) in a neighborhood, also play a crucial role, with a notable positive correlation observed. This trend suggests that younger adults are more inclined to participate in short-term rentals, thus increasing the number of listings in areas where they are concentrated. The GWR model further highlighted the spatial variability in these relationships, showing that while factors like rent prices, percentage of young adults, and number of POI were significant across the city, their influence varied spatially. This underscores the importance of local contexts and neighborhood characteristics in shaping Airbnb dynamics.

Despite its insights, this study has limitations that justify consideration. Firstly, the analysis focused solely on quantitative data, potentially overlooking qualitative aspects such as resident perceptions or Airbnb host motivations, which could provide deeper context to the spatial distribution patterns observed. Moreover, while the GWR model captured local variations, the choice of kernel bandwidth could affect the robustness of results, suggesting the need for sensitivity analysis. Additionally, the study's time frame might not capture long-term trends or seasonal fluctuations in Airbnb activity, which would require longitudinal data for a more comprehensive understanding of temporal dynamics.

For future research, incorporating qualitative methods such as interviews or surveys with residents, tourists, and Airbnb hosts could enrich the analysis by uncovering nuanced socio-cultural factors influencing Airbnb distribution. Expanding the study scope to include a broader range of cities or comparing different regulatory environments could offer insights into how local policies shape Airbnb dynamics. Furthermore, integrating real-time or dynamic data sources could enhance predictive models and facilitate more adaptive urban planning strategies. Addressing these approaches could deepen understanding and inform more effective policy interventions in managing short-term rental platforms like Airbnb.

In conclusion, this study underscores the importance of adopting a data-driven approach in understanding and managing the spatial distribution of Airbnb in urban settings like Lisbon. By integrating geographic analysis, socio-demographic insights, and advanced statistical techniques, it can help provide a framework for urban planners and policymakers to devise targeted strategies in the city of Lisbon. These strategies could include zoning regulations, infrastructure investments, and tourism management initiatives tailored to enhance the benefits of Airbnb while mitigating its potential negative impacts on local housing markets and community dynamics. Ultimately, leveraging these findings can contribute to more sustainable and equitable urban development practices in Lisbon coping with the challenges and opportunities presented by the sharing economy.

BIBLIOGRAPHICAL REFERENCES

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93-115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Antunes, G., & Ferreira, J. (2021). Short-term Rentals: How Much is Too Much - Spatial Patterns in Portugal and Lisbon. *Tourism and Hospitality Management, V*, *27*(3), 581-603. https://doi.org/10.20867/thm.27.3.6

Área Metropolitana de Lisboa. (2023, September 4). *A municipality in the Lisbon Metropolitan Area - Lisboa*. aml. https://www.aml.pt/en/municipios/lisboa/

Borges, C., Lopes, S., Fernandes, R. D. P., & Marat-Mendes, T. (2022). Planning at the edge: urbanism and socio-political transition in Chelas, Lisbon. *Planning Perspectives*, *37*(4), 761-793. https://doi.org/10.1080/02665433.2021.2001364

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. Geographical Analysis, 28(4), 281-298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Cesarani, M., & Nechita, F. (2017). Tourism and the Sharing Economy: An Evidence from Airbnb Usage in Italy and Romania. *Symphonya. Emerging Issues in Management*, (3), 32-47. https://doi.org/10.4468/2017.3.04cesarani.nechita

Cliff, A. D., & Ord, J. K. (1981). Spatial processes: Models & applications. Pion.

Esri. (n.d.). *How Geographically Weighted Regression works—ArcGIS Pro | Documentation*. ArcGIS Pro Resources | Tutorials, Documentation, Videos & More. Retrieved April 6, 2024, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-geographicallyweightedregression-works.htm

Estevens, A., Cocola-Gant, A., López-Gay, A., & Pavel, F. (2023). The role of the state in the touristification of Lisbon. *Cities*, *137*. https://doi.org/10.1016/j.cities.2023.104275

Eugenio-Martin, J. L., Cazorla-Artiles, J. M., & González-Marte, C. (2019). On the determinants of Airbnb location and its spatial distribution. *Tourism Economics*, *25*(8), 1224-1244. https://doi.org/10.1177/1354816618825415

Franco, S. F., & Santos, C. D. (2021). The impact of Airbnb on residential property values and rents: Evidence from Portugal. *Regional Science and Urban Economics*, *88*. https://doi.org/10.1016/j.regsciurbeco.2021.103667

Hilbe, J. M. (2011). Negative binomial regression (2nd ed.). Cambridge University Press.

Huxley, M., & Inch, A. (2020). *International Encyclopedia of Human Geography* (2nd ed.). Elsevier eBooks. https://doi.org/10.1016/B978-0-08-102295-5.10228-8

INE. (n.d.). *Web Portal*. Statistics Portugal - Web Portal. Retrieved April 3, 2024, from https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_sobre_o_ine&xlang=en

Inside Airbnb. (2023). *Get the Data*. Inside Airbnb. http://insideairbnb.com/get-the-data/

Ki, D., & Lee, S. (2019). Urban Planning Policies to the Renewal of Riverfront Areas: The Lisbon Metropolis Case. *sustainability*, *11*(15). https://doi.org/10.3390/su11154108

Kim, J. -J., Cho, M. -J., & Lee, M. -H. (2022). An Analysis of the Price Determinants of Multiplex Houses through Spatial Regression Analysis. *Sustainability*, *14*(12). https://doi.org/10.3390/su14127116

Kim, M. -Y., & Lee, S. -W. (2021). Regression Tree Analysis for Stream Biological Indicators Considering Spatial Autocorrelation. *International Journal of Environmental Research and Public Health*, *18*(10), 5150.

Lee, S., Won, M., & Lee, H. (2012). Testing Spatial Autocorrelation of Burn Severity. *ournal of Korean Forestry Society*, *101*(2), 203-212. https://agris.fao.org/search/en/providers/122646/records/647369b153aa8c89630dbc5c

Lee, Y.-Y., & Chang, Y. L. (2019). Uncovering Los Angeles Tourists' Patterns Using Geospatial Analysis and Supervised Machine Learning with Random Forest Predictors. *International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA*, 1275-1280. https://ieeexplore.ieee.org/document/9071300

Lestegás, I., Seixas, J., & Lois-González, R.-C. (2019). Commodifying Lisbon: A Study on the Spatial Concentration of Short-Term Rentals. *Social Sciences.*, *8*(2), 33. https://doi.org/10.3390/socsci8020033

Lisbon City Council. (2023). *Lisboa*. The economy in figures 2022. https://www.lisboa.pt/fileadmin/cidade_temas/economia_inovacao/documentos/LisboaN2022_EN.pdf

Lisbon City Council. (2023, June 15). *Município, Juntas de Freguesia - MUNICÍPIO de LISBOA*. Câmara Municipal de Lisboa. https://www.lisboa.pt/municipio/freguesias

Lisbon City Council/Georeferenced Information Management Division. (2023, 12 14). *Freguesias-2012*. Portal Dados Abertos. https://dados.cm-lisboa.pt/dataset/freguesias-2012

Louro, A., Marques da Costa, N., & Marques da Costa, E. (2021). From Livable Communities to Livable Metropolis: Challenges for Urban Mobility in Lisbon Metropolitan Area (Portugal). *International Journal of Environmental Research and Public Health*, *18*(7), 3525. https://doi.org/10.3390/ijerph18073525

Medeiros, E. (2022). Urban participatory planning approaches in capital cities: the Lisbon case. *European Planning Studies*, *30*(6), 1144-1161. https://doi.org/10.1080/09654313.2021.1973379

Medeiros, E., Brandão, A., Pinto, P. T., & Lopes, S. S. (2021). Urban Planning Policies to the Renewal of Riverfront Areas: The Lisbon Metropolis Case. *Sustainability*, *13*(10). https://doi.org/10.3390/su13105665

Nunes, M. (2023). Predicting and explaining airbnb prices in Lisbon: Machine learning Approach. *Veritati - Repositório Institucional da Universidade Católica Portuguesa*. https://repositorio.ucp.pt/bitstream/10400.14/41431/1/203253175.pdf

OpenStreetMap. (n.d.). *About*. OpenStreetMap. Retrieved May 2, 2024, from https://www.openstreetmap.org/about

Petruzzi, A. M., Carmo, M. D., & Correia, A. (2020). Airbnb and neighbourhoods: an exploratory study. *International Journal of Tourism Cities*, *6*(1), 72-89. https://doi.org/10.1108/IJTC-08-2019-0119

Pires, A., & Rafael, C. (2021). Airbnb Customer Satisfaction Through Online Reviews. In: Abreu, A., Liberato, D., González, E.A., Garcia Ojeda, J.C. (eds) Advances in Tourism, Technology and Systems. *Smart Innovation, Systems and Technologies*, *209*, 241-250.

Quattrone, G., Greatorex, A., D, Q., Capra, L., & Musolesi, M. (2018). Analyzing and predicting the spatial penetration of Airbnb in U.S. cities. *EPJ Data Science*, *7*(31). https://doi.org/10.1140/epjds/s13688-018-0156-6

Santos, T., Ramalhete, F., Julião, R. P., & Soares, N. P. (2022). Sustainable living neighbourhoods: Measuring public space quality and walking environment in Lisbon.

Geography and Sustainability, *3*(4), 289-298. https://doi.org/10.1016/j.geosus.2022.09.002

Sarkar, A., Pick, J. B., & Jabeen, S. (2024). Spatiotemporal Patterns and Socioeconomic Influences on Host Participation in Short-Term Rental Markets: Airbnb in San Francisco. *ISPRS International Journal of Geo-Information*, *13*(1). https://doi.org/10.3390/ijgi13010004

Singh, N., & Azevedo, A. B. (2021). Geography of AirBnb in Barcelona and Lisbon. *Finisterra - Revista Portuguesa de Geografia*, *56*(118). https://doi.org/10.18055/Finis23725

Sun, S., Wang, X., & Hu, M. (2022). Spatial distribution of Airbnb and its influencing factors: A case study of Suzhou, China. *Applied Geography*, *139*. https://doi.org/10.1016/j.apgeog.2022.102641

Sun, S., Zhang, S., & Wang, X. (2021). Characteristics and influencing factors of Airbnb spatial distribution in China's rapid urbanization process: A case study of Nanjing. *PLOS ONE*, *16*(3). https://doi.org/10.1371/journal.pone.0248647

Tigabu, S., Liyew, A. M., & Geremew, G. M. (2021). Modeling spatial determinates of teenage pregnancy in Ethiopia; geographically weighted regression. *BMC Women's Health*, *254*(21). https://doi.org/10.1186/s12905-021-01400-7

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240. https://doi.org/10.2307/143141

Zhang, Z., & Fu, R. J. C. (2022). Spatial Distribution of Airbnb Supply in Los Angeles. *Tourism Analysis*, *27*(4), 467-477. https://doi.org/10.3727/108354222X16571659728565

NOVA
IMS
Information
Management
School

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa