Megaron, Vol. 19, No. 3, pp. 362–374, September 2024

Megaron
https://megaron.yildiz.edu.tr - https://megaronjournal.com
DOI: https://doi.org/10.14744/megaron.2024.76735

Article

Exploring Zipf's Law and population density patterns in metropolitan İstanbul's neighborhoods: A spatial insight

Şüheda KÖSE*, Damla ERENLER

Department of City and Regional Planning, IZTECH, Izmir, Türkiye

ARTICLE INFO

Article history
Received: 24 March 2022
Revised: 05 July 2024
Accepted: 09 July 2024

Key words:
Metropolitan Istanbul; Population Density Distribution; Spatial Dependency; Zipf's Law.

ABSTRACT

Most studies on size distributions focus on examining rank-size distributions at urban or regional scale, but they often overlook their spatial dependencies, distributions, and neighboring relationships. This study aims to test Zipf's Law at the neighborhood scale of Metropolitan Istanbul, analyze its spatial dependencies, and investigate their spatial behavioral patterns in urban areas over the past decade. Initially, we found that Zipf's Law is not valid at the neighborhood scale of Istanbul. Secondly, we identified significant spatial dependencies in neighborhood population densities, observed clustering of high- and low-density neighborhoods in different locations, and detected their influences from adjacent neighborhood densities. Thirdly, we observed that population dynamics are directly affected by urban policies. Based on these findings, when spatial dependencies are considered as essential factors and analyzed in detail at lower scales, population density can provide preliminary insight into the social, economic, and political processes occurring in the city.

Cite this article as: Köse, Ş., Erenler, D. (2024). Exploring Zipf's Law and population density patterns in metropolitan İstanbul's neighborhoods: A spatial insight. Megaron, 19(3), 362-274.

INTRODUCTION

With the faster spread of neoliberal urban policies due to globalization, production outputs, capital, labor, and assets can move more easily and quickly. As a result, with the removal of physical boundaries between different geographies, social and economic interactions have developed, leading to faster and more intense urban development. Therefore, capital-oriented policies have affected spatial distribution in urban systems. In the face of all these factors, urban systems, as noted by Dicken & Lloyd (1990), may exhibit regular distributions in terms of density and size. One of the theories concerned with the regular distribution of population sizes in settlements is Zipf's Law, also known as the rank-size rule (Gabaix, 1999).

Zipf's Law was developed by linguist George Kingsley Zipf in 1949 (Zipf, 1949). The law posits an inverse relationship between the frequency of values subjected to ranking and their rankings (Nitsch, 2005; Zipf, 1949). Although Zipf developed his theory in the field of linguistics, it has been tested in various fields, from economics to the arts, such as company sizes, income distribution, industrial sectors, web pages, book analyses, and song lyrics. It has also been applied and tested in the field of urban planning to analyze the population distributions of cities. According to urban

***Corresponding author**
*E-mail adres: suhedakose@iyte.edu.tr

Published by Yıldız Technical University, İstanbul, Türkiye

researchers (Arshad et al., 2018), a city's optimal distribution should adhere to Zipf's Law. In other words, the hierarchy between cities should exhibit an inverse proportionality between their population rankings and frequencies.

In the 1950s, Zipf's Law began to be discussed in the field of urban and regional planning. Academic studies testing the law between cities and regions showed its validity (Corominas-Murtra & Solé, 2010; Dicken & Lloyd, 1990) and provided some preliminary insights into analyzing the socio-economic levels of cities (Córdoba, 2008). The impact of globalization, which began in the 1970s, has led to differences in studies on Zipf's Law. Studies conducted by urban scholars such as Akseki et al. (2014), Behrens et al. (2014), Black & Henderson (2003), and Casetti (1972) found the loss of validity of Zipf's Law in cities and regions with rapid urbanization rates.

With metropolitan cities becoming the central hubs of knowledge flow in the 2000s (Mukherji & Silberman, 2018), Zipf's Law started to vary according to cities and regions. Hackmann & Klarl (2020) found that the relevant law is more valid among medium-sized cities than mega-cities, while Giesen & Südekum (2011) analyzed its validity among cities in developed regions. Sun et al. (2021) observed a more unequal population distribution in cities with advanced industrial structures, whereas Luckstead & Devadoss (2014) found the law to be invalid even among the world's largest cities. A study testing Zipf's Law at the neighborhood scale of 12 global cities found that neighborhood sizes conform to the law (Sahasranaman & Jensen, 2020). Therefore, although population distributions vary according to scale, there is a direct relationship between the urbanization rate of cities and the consistency of the law (Kundak & Dökmeci, 2018).

Istanbul is the most rapidly urbanizing metropolitan city in Türkiye. Especially in the last 50 years, it has been the city with the highest population density due to irregular domestic and international migration rates (Bayartan, 2003). However, according to the data from the Turkish Statistical Institute (TSI), the increasing migration rate started to decline after 2015 (TSI, 2024). The decline in migration rates has also led to a decrease in population size. Istanbul, the fastest-urbanizing city in Türkiye, was chosen as the case study to examine population changes over the past decade in detail. To access reliable results, both the spatial and numerical distributions of irregular population movements have been examined for annual changes.

Based on the aforementioned reasons, this study aims to test the validity of Zipf's Law through Istanbul's neighborhood settlements, examine spatial dependencies, and identify the neighborhoods where population density varies locationally. In line with the stated objectives, the following research questions will be addressed:

- Is Zipf's Law valid at the neighborhood scale in Istanbul? If so, how do neighborhood population densities exhibit spatial distribution patterns?
- Is there spatial dependence in the distribution of neighborhood population densities in Istanbul? If so, how strong of a factor might this be?

This research provides answers to the above questions through the Istanbul case, contributing to the literature both methodologically and theoretically. Methodologically, in addition to Zipf's Law, physical distance has been used as an indicator. This indicator has been tested using spatial econometric methods, leading to the development of a multiple hybrid approach model. Theoretically, it has been demonstrated that the law can be tested through the distribution of population density, which is rarely used in the literature. It also emphasizes the need to include spatiality in population analysis studies.

This article is structured as follows: Section 2 covers the development of Zipf's Law, its use in urban studies, and the findings obtained. Sections 3 and 4 provide detailed information on the study area, the dataset, and the methods used in the research. Section 5 analyzes the findings of the study, while Section 6 contains a general evaluation of the findings.

Literature Review

Studies on the size distributions of cities are based on two main theories: Central Place Theory and Zipf's Law. Central Place Theory conceptualizes the relationships between settlements and focuses on the gradual organization of settlements. Developed by Christaller (Brush, 1966; Christaller, 1966) and Lösch (1954), Central Place Theory predicts that urban systems will develop in a functional hierarchical structure based on the goods and services they provide. According to this theory, there is a close relationship between the quantity of goods and services provided by a city at certain intervals, its spatial market influence, population density, and the demand rate of this population (LeSage, 1999). Therefore, settlements in a country are expected to be distributed in certain sizes and geographical distributions depending on these factors.

According to Beckmann (1999), urban systems are also shaped by political decisions. In this context, if cities are divided into size groups, the distributions of their numbers and the areas they cover can be predicted or ranked in a regular manner. Over time, researchers such as Lösch (1954), Beckmann (1999), and many others have developed the thesis that Central Place Theory and Zipf's Law are compatible or shaped by the same phenomena, but as noted by Parr (1985), no definitive and widely accepted reconciliation has been reached between the two models.

Unlike the hierarchical organization of central places, Zipf's Law focuses on the growth of cities at national, regional, or urban scales (Zipf, 1949). Zipf's Law states that there is an inverse relationship between the ranking of population sizes of settlements (regions, cities, towns, neighborhoods, etc.) and the populations of these settlements at defined time intervals (Parr & Suzuki, 1973). The rank-size rule is also described as a negative relationship between the logarithms of urban populations and the logarithms of the ranks of population sizes. In this case, the slope of the curve that represents the relationship between the logarithm of population size and the logarithm of rank size is equal to -1 (Knudsen, 2001). This idea is based on the assumption of a regular relationship between the populations and rankings of settlements.

Zipf's Law, one of the settlement hierarchy theories, was first proposed by Auerbach in 1913 (Auerbach, 1913). The law was shaped by Auerbach's Pareto coefficient (Ioannides & Overman, 2003; Nitsch, 2005). The Pareto coefficient indicates how evenly the population is distributed among cities. When the Pareto coefficient is greater than 1, small settlements have relatively high proportions of the population, and population distribution is concentrated in these areas. When it is less than 1, the population is more concentrated in large cities (Marin, 2007). Over time, changes in intercity relationships occur with the rapid growth of large cities and the stagnation of small cities. Casetti (1972), who developed the extended rank-size rule, aimed to test changes in coefficients over time, so he reformulated Zipf's model by adding a time coefficient (Dokmeci & Turk, 2001).

Research on Zipf's Law has been reformulated and developed over time. However, the relationship between spatial dependence and the law in rank-size distribution studies has been discussed for the past three years (Bergs, 2021). The theory of spatial dependence stems from the First Law of Geography. According to Tobler, "*everything is related to everything else, but near things are more related than distant things*" (Tobler, 1970:3). Building upon this law, spatial dependence is defined as the degree of spatial autocorrelation between independently measured values in a geographical area (Kitchin & Thrift, 2009). For example, it assumes that there is autocorrelation, or a relationship, between a measured value in a geographic unit and the same type of value in its neighboring unit (Anselin, 1985). Therefore, while Zipf's rank-size distribution hierarchy is measured by population sizes, these population sizes are not spatially independent; they functionally exhibit autocorrelation. Hence, when measuring rank-size distributions, the level of functional relationship between the values' physical locations also needs to be tested.

When examining studies on the distribution of city sizes in Türkiye, the linearity of the city size distribution is generally assumed in the literature as it stands. However, there is no common consensus among the results obtained from these studies. Dokmeci (1986) applied the rank-size rule both at the national and regional levels in Türkiye between 1945 and 1975. She found the rank-size rule to be invalid due to the shaping of the size distribution of regions in parallel with their economic development since 1945.

Between 1975 and 1982, the distribution of city sizes in Türkiye showed a better fit to the rank-size rule. Marin (2007) examined the population changes between cities in 1985, 1990, and 2000 by using econometric methods. It was found that the Pareto coefficient was below -1, indicating a departure from Zipf's Law in urban population distribution in Türkiye after 1985 (Marin, 2007). In a study conducted by Deliktaş et al. (2013), it was found that the Pareto exponent of 81 provinces in Türkiye varied between 0.87 and 0.97 during the period of 1980-1997, indicating a more linear spread of sizes.

In the 2000s, studies revealed imbalances in rank-size distribution. A study in 2015 concluded that Zipf's Law failed among cities in Türkiye (Duran & Özkan, 2015). In another detailed study, data obtained from 973 districts in Türkiye were used, and it was found that the sizes of districts were unevenly distributed in 2019, indicating the invalidity of the rank-size rule. (Duran & Cieślik, 2021).

In conclusion, Zipf's Law, subjected to testing through examples from Türkiye and around the world, has not found its full counterpart in urban systems. There are three main reasons for this. Firstly, research has been extensively conducted at the national, regional, and city scales, but rare studies have been encountered at the district and neighborhood scales. Secondly, the level of development in many countries, regions, and cities has led to differences in size distribution. Thirdly, when testing Zipf's Law, the locations with high population densities have been neglected. Therefore, this study, which is distinct from other studies, evaluates Zipf's Law at the neighborhood scale. It questions the validity of the law through population density distribution. While elucidating Zipf's Law with a spatial interaction, the research also provides an opportunity to measure the level of spatial impact in detail through spatial econometrics.

MATERIALS: STUDY AREA

According to the Turkish Statistical Institute (TSI, 2021), Istanbul, with its approximately 16 million population, is the most populous city in Türkiye and the largest metropolitan area in the country, offering various urban services such as economic, social, cultural, historical, and transportation services. When examined in terms of population density, the city receives an average of around 29% internal and 7% external migration annually (TSI, 2021). Despite Istanbul's high population density, uneven population growth rates and migration rates have been observed in the past decade. For this reason, it has been selected as the study area to examine the changes in density distribution at the neighborhood scale over the years.

As a preliminary data analysis of the study, the spatial sizes of districts and neighborhoods in Istanbul are shown in Figure 1. Istanbul, which has 39 districts, consists of 963 neighborhoods as of the year 2020. Şile, located in the easternmost part of the city on the Anatolian side, has the highest number of neighborhoods with 62, while Adalar, located in the south of the city, has the lowest number of neighborhoods with 5 (TSI, 2021).

Figure 2 illustrates the spatial distribution of neighborhood population densities in Istanbul for the years 2010 and 2020. The neighborhoods with the highest population density are concentrated in the southern districts of the city. In 2010, a total of 12 neighborhoods in Esenler, Bağcılar, Kağıthane, and Esenyurt districts, and in 2020, a total of 21 neighborhoods in Bağcılar, Güngören, Kağıthane, Esenyurt, and Zeytinburnu districts had the highest population density with 347 people per hectare.

The neighborhoods with the lowest population density are located on the peripheries on both sides of the city. Şile and Çatalca districts have the lowest population density with 3 people per hectare in a total of 23 neighborhoods. Medium-sized neighborhoods with population densities ranging from 141 to 280 people per hectare are distributed around high-density neighborhoods. Neighborhoods with population densities ranging from 71 to 140 people per hectare are distributed around low-density neighborhoods. Additionally, when compared between the two years, the scattered population density of neighborhoods in 2010 exhibited a more concentric spatial distribution in the districts and neighborhoods with the highest population density in 2020.

Table 1 displays the summary of changing population data in Istanbul from 2010 to 2020. Clearly, the average neighborhood size decreased from 16,750 in 2010 to 16,056 in 2020. The largest neighborhood's size expanded from 84,560 in 2010 to 101,660 in 2020. Moreover, the number of neighborhoods with a population of 15,000 or more increased from 364 in 2010 to 417 in 2020. At first glance, there appeared to be a balanced distribution among neighborhood sizes and numbers. However, upon closer examination of minimum and maximum sizes at any given time, there is not a consistent and stable distribution among neighborhoods. Additionally, it is observed that the minimum and maximum neighborhood size distributions began to change in 2015 and 2016.

When the population data in Table 1 is examined in

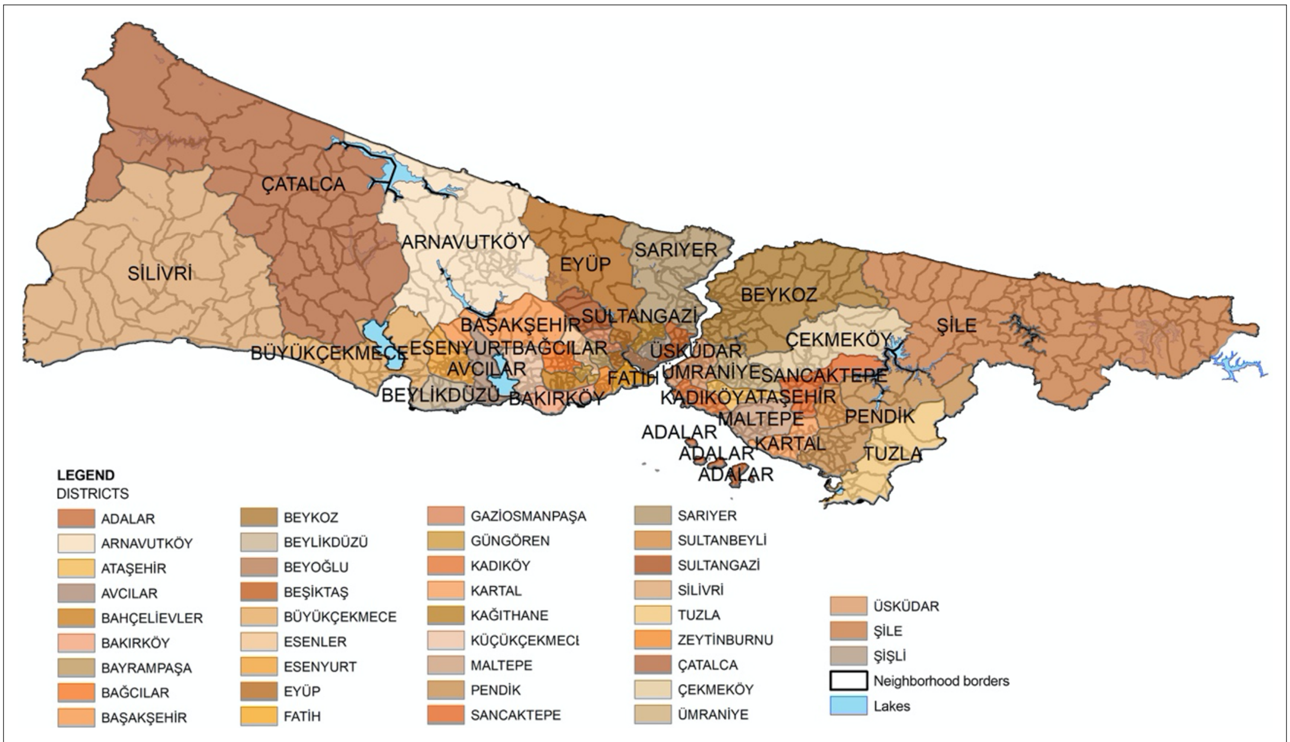

Figure 1. Study Area.

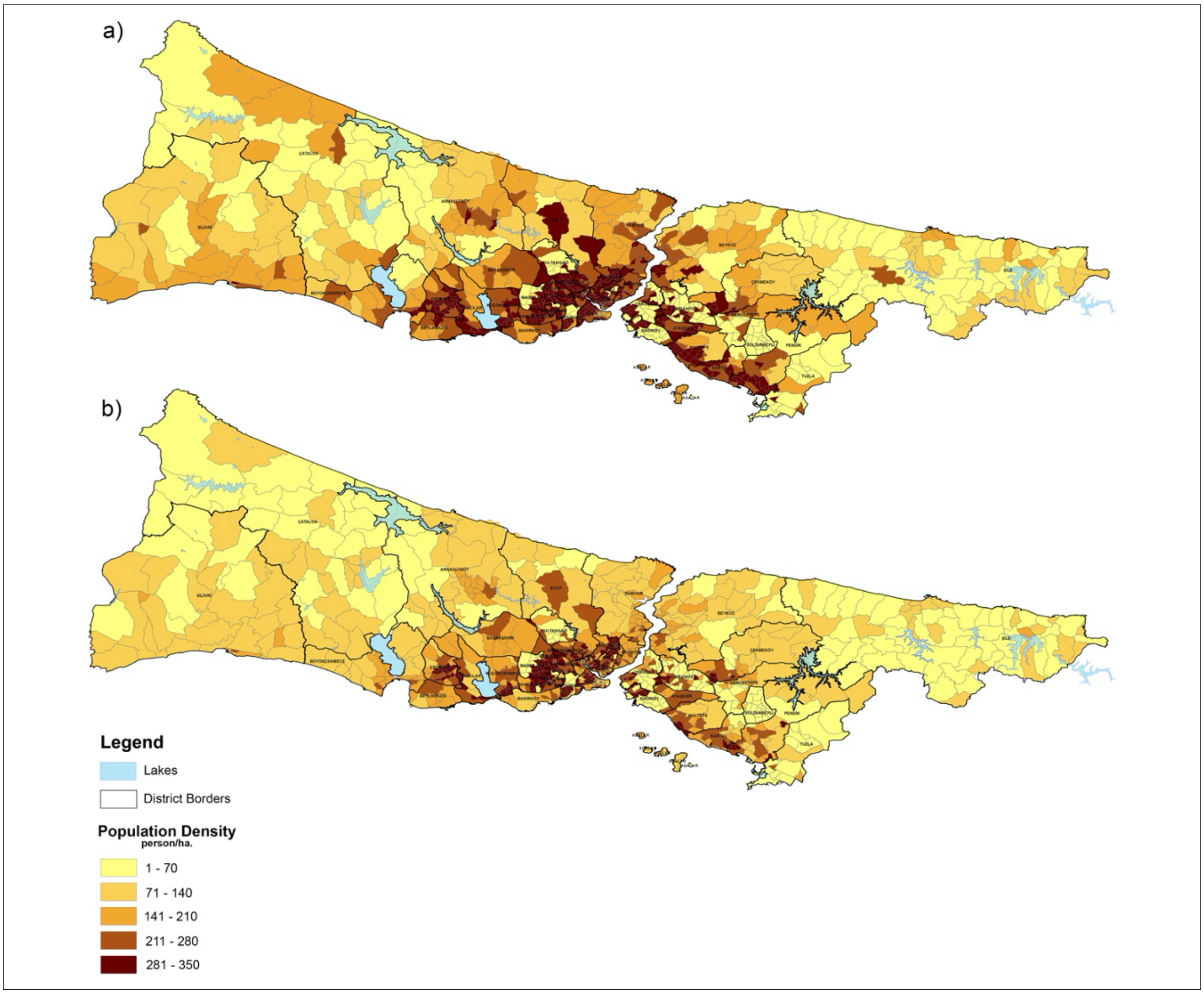

Figure 2. **(a)** Population density distributions of neighborhoods in Istanbul in 2010, **(b)** Population density distributions of neighborhoods in Istanbul in 2020.

detail, two significant years affecting population change stand out: 2013 and 2020. The first population change was driven by the population increase rates, migration rates, and the number of neighborhoods in 2013 and 2020. The reason for the initial population change was the transformation of villages into neighborhoods under metropolitan municipalities through Law No. 6360 issued in the last month of 2012 (TC Resmi Gazete, 2012), coupled with the influx of Syrian refugees into the country starting at the end of 2011. This law introduced a different dimension to the urbanization process in the country, leading to a rapid increase in the urban population. The population of Istanbul, which was 13,710,512 in 2012, reached 14,160,467 in 2013 with a population growth rate of 3.28% (TSI, 2024). With the opening of doors to Syrian refugees, the foreign population, which was 85,360 in 2012, increased by approximately 50,000 people in 2013, reaching 135,018 (TSI, 2024). Consequently, 2013 had the highest population growth rate within the study period.

The second population change occurred in 2019 due to the global pandemic outbreak. Istanbul was negatively affected by the COVID-19 pandemic, which claimed the lives of approximately 10,000 people worldwide per month (WHO, 2024). The population growth rate and migration rates, which had been increasing until 2019, decreased significantly in 2020 for the first time. While the population growth rate decreased by 3.35% and the migration rate decreased by 5.5%, the population density remained constant. In summary, population changes, as observed in the Istanbul case, can be influenced by economic, social, and political decisions (Sun et al., 2021).

Table 1. Descriptive Summary Statistics of Neighborhoods

Year	2010	2011	2012	2013	2014	
Avg. neighborhood size (1000 people)	16.75	17.21	17.48	15.12	15.34	
Median neighborhood size (1000 people)	13.93	14.17	14.58	11.79	11.93	
Minimum neighborhood size (1000 people)	0.018	0.015	0.014	0.022	0.013	
The neighborhood and its district with the smallest size	Sarıdemir Neigh. (Fatih)	Sarıdemir Neigh. (Fatih)	Sarıdemir Neigh. (Fatih)	Sarıdemir Neigh. (Fatih)	Sarıdemir Neigh. (Fatih)	
Maximum neighborhood size (1000 people)	84.56	86.43	85.65	86.63	85.75	
Neighborhood and district with the largest size	Zafer Neigh. (B.evler)	Zafer Neigh. (B.evler)	Zafer Neigh. (B.evler)	Zafer Neigh. (B.evler)	Zafer Neigh. (B.evler)	
Population density (person/ha)	25	26	26	27	28	
Population growth rate (%)	2.64	2.76	1.68	3.28	1.52	
Net migration rate (%)	16.53	16.68	13.92	15.49	15.39	
Total number of neighborhoods	783	783	783	936	937	
Number of Neighborhoods> 15,000	364	370	376	382	389	
Year	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Avg. neighborhood size (1000 people)	15.26	15.42	15.65	15.64	16.11	16.05
Median neighborhood size (1000 people)	12.18	15.15	12.18	12.17	12.48	12.39
Minimum neighborhood size (1000 people)	0.011	0.013	0.012	0.019	0.014	0.013
The neighborhood and its district with the smallest size	OSB Neigh. (İkitelli)	Sarıdemir Neigh. (Fatih)	Sarıdemir Neigh. (Fatih)	OSB Neigh. (Dudullu)	OSB Neigh. (Dudullu)	OSB Neigh. (Dudullu)
Maximum neighborhood size (1000 people)	86.44	88.95	93.22	95.63	98.80	101.66
Neighborhood and district with the largest size	Zafer Neigh. (B.evler)	Atakent Neigh. (Halkalı)	Atakent Neigh. (Halkalı)	Atakent Neigh. (Halkalı)	Atakent Neigh. (Halkalı)	Adnan Kahveci Neigh. (Beylikdüzü)
Population density (person/ha)	28	28	29	28	29	29
Population growth rate (%)	1.95	1.00	1.52	0.25	2.99	-0.36
Total migration rate (%)	15.72	12.62	13.96	12.98	15.98	10.48
Total number of neighborhoods	960	960	960	963	963	963
Number of Neighborhoods> 15,000	402	412	418	412	421	417

Data Set and Methodology

The study encompasses all neighborhoods in Istanbul for the years 2010 and 2020. The neighborhood population data were obtained from the Turkish Statistical Institute in 2021 (TSI, 2021). The population densities of neighborhoods in 2010 and 2020 were comparatively analyzed using established methods.

After reviewing the literature, it is evident that Zipf's Law examines the rank-size relationships of cities and regions, Exploratory Spatial Data Analysis (ESDA) evaluates their spatial relationships, and Spatial Autoregressive (SAR) and Spatial Error (SEM) test spatial dependency. Typically, researchers have tested size distributions at the city or regional scale using general assumptions. However, there are a few studies that investigate density distributions, examine relationships between densities from the part to the whole scale (neighborhood to city), and assess their spatial interdependence. Therefore, this study presents a comprehensive hybrid model.

The research methodology consists of three parts. In the first stage, the hierarchy of neighborhood population densities was measured using Zipf's Law. In the second stage, the measurement results were tested for spatial dependence using the Spatial Error Model (SEM) and Spatial Autoregressive Model (SAR) and were compared with the findings of Zipf's Law. In the third stage, spatial autocorrelation between neighborhood locations and population densities was tested using Global Moran's I, and neighborhood adjacency relationships were analyzed using LISA analysis.

Zipf's Law

Zipf's Law, developed by George K. Zipf (1949), is a practical method that has been used for many years to test the growth patterns of cities, thereby analyzing the acceptability of social and economic growth theories (Brakman et al., 1999). Zipf's Law assumes that the distribution of neighborhoods should be linear according to the rank-size relationship. If a city's Zipf distribution is valid, it implies that its growth is sequential, orderly, and controlled. Formula 1 illustrates Zipf's Law. In the formula, N represents the sample size, A represents the constant empirically obtained from the data, x denotes the rank of population density in a neighborhood, Prob(x) represents the probability of population density in a neighborhood at rank x, and Freq(x) denotes the frequency of the neighborhood at rank x (Gabaix & Ioannides, 2004).

$$\log(\chi) * \log(\text{freq}(\chi)) = A * N \qquad \text{Formula (1)}$$

If we rank neighborhoods according to population density, we observe log(x) on the x-axis and freq(x) on the y-axis (Knudsen, 2001). If the slope is greater than 1, it indicates that neighborhoods are dispersed. If it is smaller than 1, it suggests that they are more concentrically clustered. If a straight line with a slope of -1 is observed, it indicates the validity of Zipf's Law, demonstrating that neighborhood densities exhibit equal or similar distributions. However, when testing neighborhood densities, the law does not account for the relationships between neighborhoods.

Spatial Error Model (SEM) and Spatial Autoregressive Model (SAR)

In a multiple linear regression model ($\gamma = \chi\beta + \varepsilon$), where y is the dependent variable, x is the explanatory variable, β is the regression coefficient, and ε is the error term (Fischer & Wang, 2011). The model assumes that the error terms are independent of each other, have a zero mean, and have a constant variance with a normal distribution. However, if the errors contain spatial autocorrelation, this assumption loses its validity (Anselin & Rey, 1991).

According to Anselin (1988), spatial dependence refers to the existence of a functional relationship between events occurring at specific points in space and those in other regions. In other words, the value of a variable at locations a and b is explained by both internal values (the internal conditions of a and b) and external values (the influence of a on b, and the influence of b on a) (Zeren, 2010). Ignoring spatial dependence can lead to underestimated true variance and problems in the estimation and interpretation of results.

Spatial regression models are determined based on the cause of spatial autocorrelation in two ways (SEM and/or SAR). SEM assumes that spatial autocorrelation exists among the error terms (Zeren, 2010; Anselin, 1988). SAR, on the other hand, assumes that the variables in neighboring locations also affect the dependent variable of observations in that location (Fischer & Wang, 2011). In other words, the model expresses the relationship between the dependent variable measured in one location and another. In the SAR method, spatial autocorrelation exists among the dependent variables.

$$\gamma = \chi\beta + \varepsilon,\ \varepsilon = \lambda W_2 e + u,\ u \sim N(0, \sigma^2 I_n) \qquad \text{Formula (2)}$$

Formula 2 represents the spatial error model. Here, γ represents the dependent variable of size nx1, X denotes the independent variable matrix of size kxk, β is the coefficient vector of size kx1, and ε represents the vector of independently and identically distributed error terms of size nx1. The spatial error coefficient, λ, measures the degree of spatial dependence among the error terms, and this coefficient takes values less than 1 (Zeren, 2010). A significant spatial coefficient indicates spatial dependence among the error terms.

$$\gamma = \rho W y + \beta x + u \text{ and } u \sim N(0, \sigma^2 I_n) \qquad \text{Formula (3)}$$

The spatial lag model is shown in Formula 3. Here, γ represents the dependent variable of size nx1, X denotes the explanatory variable matrix of size nxk, β is the coefficient vector of size kx1, and u represents the error term. The coefficient ρ, which is the dependent variable of the spatial lag model, measures the effect of the γ's in neighboring locations on the γ in the respective location. Generally, $|\rho|<1$ is assumed (Fischer & Wang, 2011). A significant ρ value indicates the presence of spatial lag dependence and implies that the use of the classical regression model is not appropriate (Çetin & Sevüktekin, 2016).

Exploratory Spatial Data Analysis (ESDA)

Exploratory Spatial Data Analysis is one of the most used methods to test the presence of spatial autocorrelation in cities and regions. This method conducts two different spatial measurements: global (Global Moran's I) and local spatial autocorrelation (Local Moran's I). Global Moran's I assumes that all observations within the studied city are connected, and a change in one neighbor affects all neighbors (Anselin, 1995). The Global Moran's I statistic is expressed as follows (Rey & Montouri, 1999):

$$I_t = \left(\frac{n}{S_o}\right) \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{i,j}, X_{i,t}, X_{j,t}}{\sum_{i=1}^{n} \sum_{j=1}^{n} X_{i,t}, X_{j,t}} \qquad \text{Formula (4)}$$

Formula 4 represents i and j as neighbors, n as the number of neighbors, $X_{i,t}$ as the density of a neighborhood in year t, W_{ij} as the standardized spatial weight matrix, and So as the sum of all W_{ij} values. If i and j share a spatial neighborhood, W_{ij} takes the value of 1; otherwise, it takes the value of 0 (Rey & Montouri, 1999). Although the Global Moran's I test provides a general discourse on the change in neighborhood densities, it does not analyze the locational information of the change.

To test the locational change in neighborhood densities, the Local Moran's I method (LISA) is used. The Local Moran's I (LISA) evaluates whether neighborhood densities differ from surrounding neighborhoods using location data. Additionally, it provides an opportunity to analyze densities in detail according to their spatial relationships. The Local Moran's I is shown as follows:

$$I_t = \left(\frac{X_i}{S_o} \right) \sum_{i=1}^{n} \sum_{j=1}^{n} W_{i,j}, X_{i,t} \text{ with } m_0 = \sum_{i}^{n} x_{i,t}^{2}$$

Formula (5)

In contrast to Global Moran's I, this method includes the m_0 value. m_0 represents the sum of the weighted spatial matrix, i.e., the sum of the W_{ij} elements (Rey & Montouri, 1999).

RESULTS AND DISCUSSION

a. Zipf's Law and Spatial Dependency Test Results

Zipf's Law was applied to examine the population density distributions in neighborhoods. While Zipf's Law provides insight into the rank-size relationship, it does not offer an explanatory method for the spatial dependency and locations of the ranks. Therefore, spatial lag and spatial error models were utilized to test the validity of Zipf's Law and spatial dependency.

Table 2 represents the results of Zipf's Law and spatial dependency tests for the years 2010 and 2020. In terms of neighborhood population densities, Zipf's Law is not valid for both years. When evaluating the spatial dependencies of population densities, it was found that their distributions are statistically positive, indicating that neighborhoods of similar sizes tend to cluster in certain urban areas. In other words, over the decade, although the rank-size distribution of Istanbul may not be valid, the population densities within each year tend to cluster in different sizes within specific regions.

Figure 3 displays the graphs of Zipf's Law, where the straight line represents the logarithm of neighborhood rankings on the y-axis to the logarithm of neighborhood population densities on the x-axis. According to the graph for 2010, the distribution in neighborhoods with high densities tends towards linearity, whereas it deviates from linearity towards neighborhoods with lower densities. The graph for 2020 exhibits a similar trend to that of 2010. However, unlike the previous measurement, in 2020, the population increase has concentrated between ln(3) and ln(5), and also at ln(12). In summary, it can be stated that Zipf's Law is not valid for both graphs, and the distributions among neighborhoods are significantly far from linear due to differences in density. Additionally, the requirements of the law explain the impact of the dependent variable on the independent variable by 54% in 2010 and 41% in 2020. The explanatory rate of the variables and the resulting numerical values (-0.0098 and -0.0039) indicate that neighborhoods in Istanbul will not be distributed in accordance with the rank-size relationship and will continue to grow through clustering.

Zipf's Law and spatial dependency tests are useful in testing rank-size relationships and spatial dependencies among

Table 2. Results of Zipf's Law and Spatial Dependency Tests in Istanbul

	Model 1 (2010)			**Model 2 (2020)**		
	Coefficient		**p**	**Coefficient**		**p**
Constant	646.589		0.0000	315.435		0.0000
Pop_Density	-0.0098		0.0000	-0.0039		0.0000
R-Squared		0.54			0.41	
Spatial Error Model						
Constant	602.869		0.0000	312.545		0.0000
Pop_Density	-0.0074		0.0000	-0.0038		0.0000
LAMBDA	0.5942		0.0000	0.1201		0.0155
R-Squared		0.66			0.42	
Spatial Lag Model						
Constant	375.488		0.0000	275.583		0.0000
Pop_Density	-0.0068		0.0000	-0.0037		0.0000
W_Rank	0.452635		0.0000	0.1400		0.0026
R-Squared	0.66		0.42			
N		783			963	

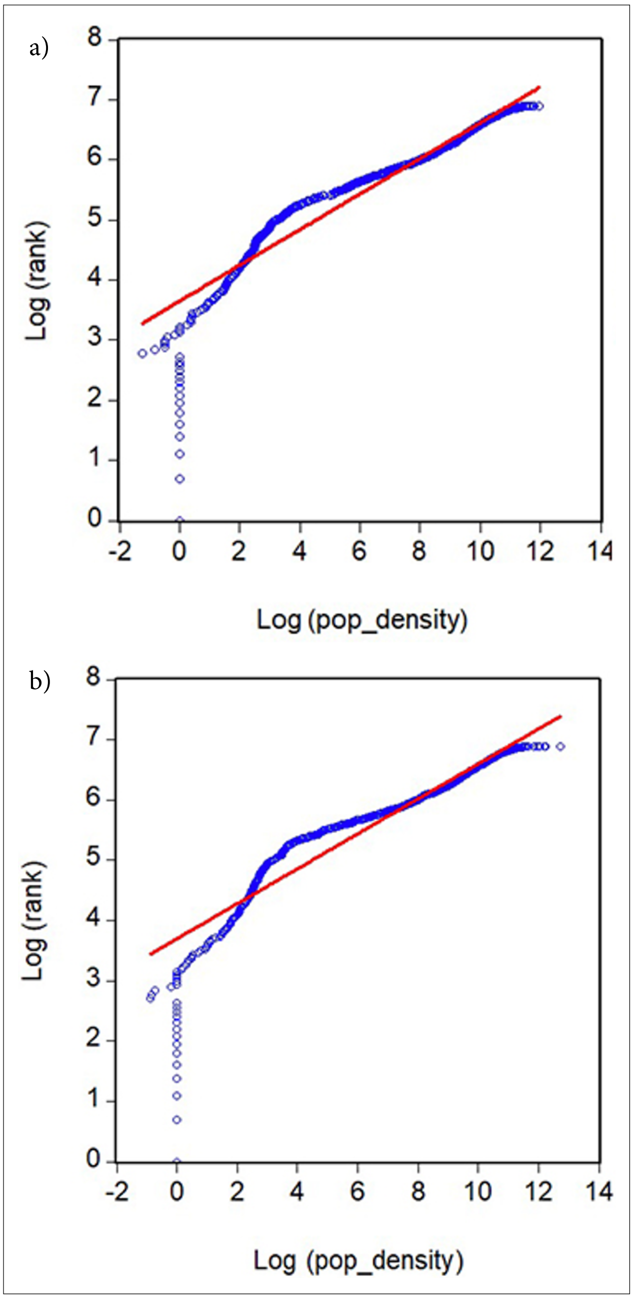

Figure 3. **(a)** Graph of Zipf distribution in 2010, **(b)** Graph of Zipf distribution in 2020.

neighborhood population densities but do not provide detailed information about their locations and neighborhood relationships. Therefore, Global Moran's I and LISA analyses were conducted to examine the correlation between neighborhood locations and population densities in detail.

b. Global Moran's I

To determine the relationships between neighborhood population densities and their locations, the Global Moran's I test was conducted. Global Moran's I explains a relationship or clustering between neighborhood density values and their locations within the range of -1 and 1. As shown in Formula 4, if the result is close to 1, it indicates positive autocorrelation, and if it is close to -1, it indicates negative autocorrelation.

Figure 4 illustrates the Global Moran's I scatter plots of neighborhood population densities for the years 2010 and 2020. The plots indicate that the density values are not randomly distributed. The Moran's I value, calculated as 0.506 in 2010, decreased to 0.364 in 2020. These values indicate that neighborhood population densities exhibit positive spatial autocorrelation. In other words, there is a positive relationship between neighborhood locations and their densities in both years. However, the positive autocorrelation value of the spatial pattern for Istanbul in 2010 is higher compared to 2020. Therefore, although neighborhood population densities have positive autocorrelation in recent years, the relationship between neighborhood locations and their densities appears to be decreasing.

c. Local Moran's I

To examine the spatial behavior between neighborhood population densities and their neighborhood relationships, Local Moran's I (LISA) analysis was utilized. As shown in Formula 5, Local Moran's I analysis tests local-scale spatial neighborhood relationships of the Global Moran's I test in four different types. HH and LL regions represent neighborhoods with positive clustering, while HL and LH regions represent neighborhoods with negative clustering (spatial outliers).

Figure 5 illustrates the local distributions of neighborhood population densities for the years 2010 and 2020. Neighborhoods located in the HH region exhibit positive clustering relationships in the southern part of the European side of the city, with their population densities higher than the Istanbul average. In 2010, the positive clustering pattern of neighborhoods with high densities extended to influence 11 neighborhoods in 2020, thereby increasing the level of spatial autocorrelation. Neighborhoods in the LL region demonstrate positive clustering along the east-west axis on both sides of the city. These neighborhoods have significantly lower population densities than the Istanbul average and are also in proximity to neighborhoods with low densities. Comparing the LL regions of the two years, the number of neighborhoods exhibiting positive clustering patterns increased in 2020, including 9 additional neighborhoods. This increase, in 2010, encompassed neighborhoods previously exhibiting negative autocorrelation, clustering in the inner areas of both sides of the city.

Neighborhoods in the HL region exhibit negative autocorrelation. The population densities of neighborhoods in these regions are above the Istanbul average. However, most neighborhoods in the HL region are adjacent to those

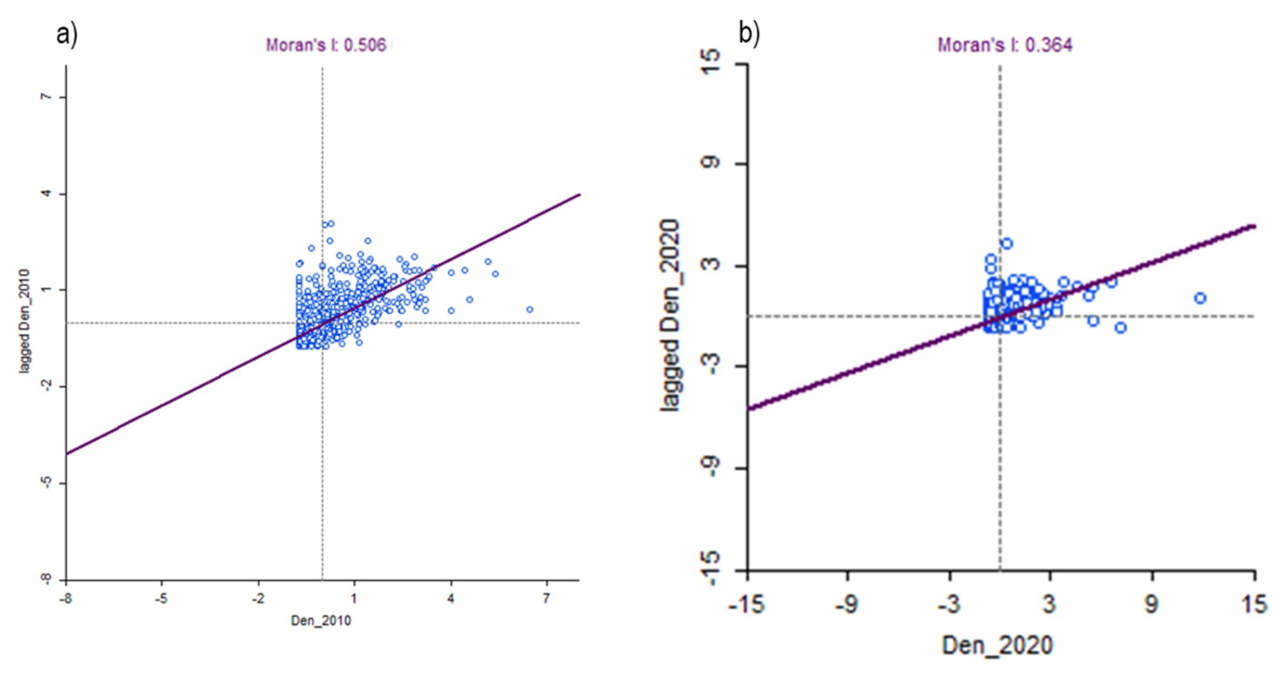

Figure 4. **(a)** Global Moran's I Distribution in 2010, **(b)** Global Moran's I Distribution in 2020.

in the LL region. Neighborhoods in the LH region represent a negative clustering pattern, with population densities lower than those in Istanbul on average. Additionally, these neighborhoods are related to those with denser populations than the Istanbul average. The number of LH regions decreased from 2010 to 2020, reducing the level of negative autocorrelation. As expected, these are observed on the European side of the city, where the HH region clusters densely.

The study examines the spatial behavior of urban growth in Istanbul, Türkiye's fastest-growing city in the last decade, both through spatial econometric tests and Zipf's Law at the neighborhood level. To conclude, Zipf's Law was invalid in both years, and there were significant differences in neighborhood population density distributions. Additionally, neighborhood densities exhibited noteworthy spatial dependence, positive spatial autocorrelation, and physical neighborhood relationships. In a general evaluation, although Zipf's Law was found to be invalid in Istanbul's density distribution, as Anselin (1995) pointed out, the spatial relationships of neighborhoods were found to be significantly affected by internal and external factors.

EVALUATION AND CONCLUSION

The aim of the study was to test the distribution of neighborhood densities using Zipf's Law and to examine their spatial dependence, clustering patterns, and neighborhood relationships in Istanbul. To achieve this goal, a multifaceted methodological approach was adopted, and both population density distributions and their spatial relationships were tested with different methods, and the validity of the acquired knowledge was verified.

Firstly, the validity of Zipf's Law and spatial dependence were measured across the city. In Istanbul, Zipf's Law proved invalid, yet it showed positive spatial autocorrelation. The dependency tests received statistically significant findings, demonstrating that neighborhood population densities were influenced by the densities of adjacent neighborhoods. It was observed that political decisions led to the invalidation of Zipf's Law in Istanbul. The "Metropolitan Law" enacted in 2012 and the influx of Syrian refugees in 2013 introduced a different perspective to urbanization in Istanbul. As mentioned in the literature, uncontrolled urbanization has had a negative impact on the maintenance of Zipf's scientific law (Dittmar, 2009). The findings regarding Zipf's distribution in Istanbul align with research conducted by Akseki et al. (2014), Behrens et al. (2014), Black & Henderson (2003), and Casetti (1972).

Secondly, the spatial behavior of density distributions was examined using Global Moran's I and Local Moran's I analyses. According to the Global Moran's I test, neighborhood population densities in Istanbul exhibited positive spatial autocorrelation. Local Moran's I analysis revealed that neighborhood population densities in Istanbul had a heterogeneous distribution. High-density neighborhoods clustered in the inner areas of both sides of Istanbul, while low-density neighborhoods were observed to have neighborhood relationships on the eastern and

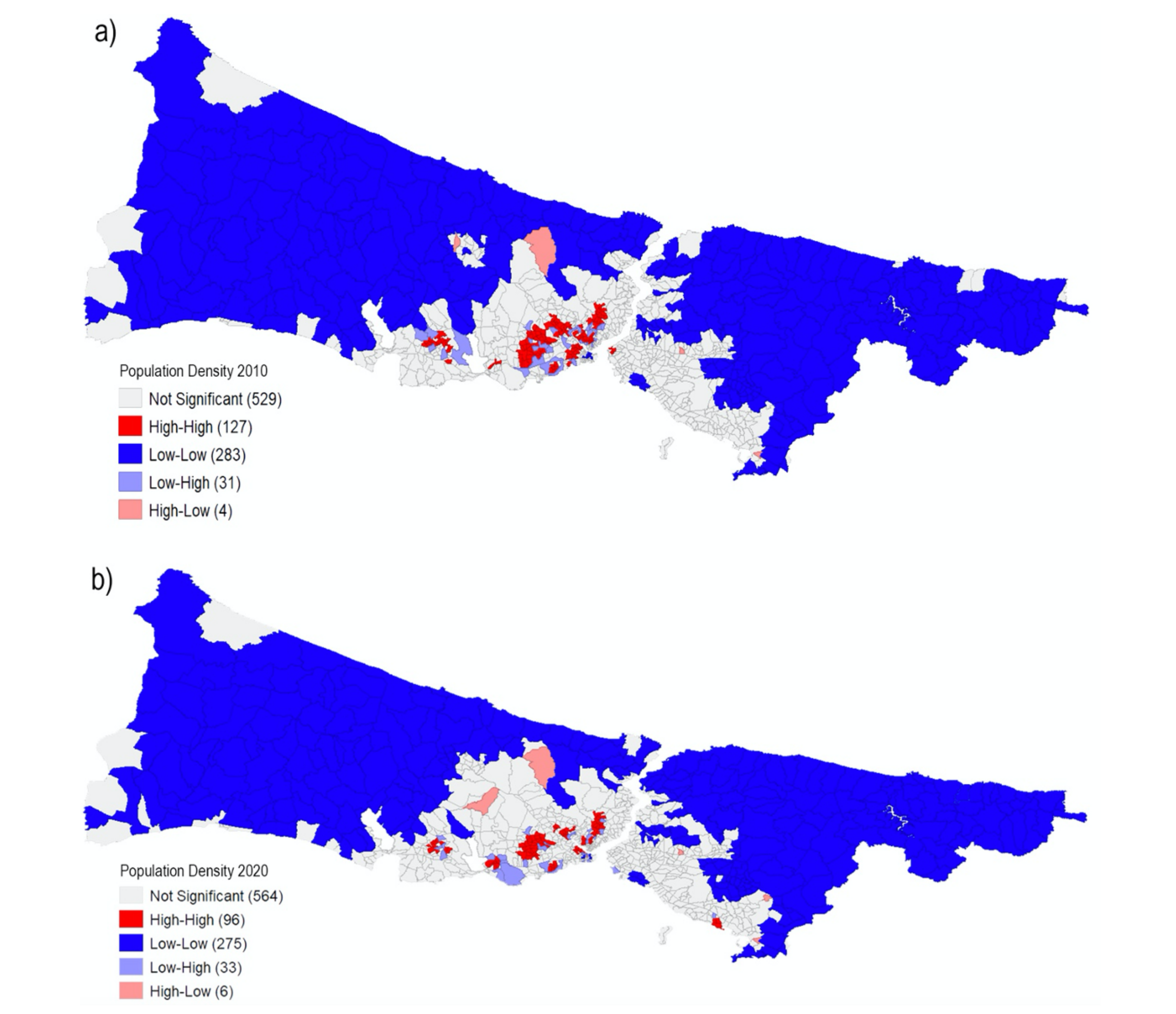

Figure 5. **(a)** Local Moran's I Distribution in 2010, **(b)** Local Moran's I Distribution in 2020.

western peripheries of the city.

In summary, it was analyzed that neighborhoods are not independent of their neighbors and are influenced by their density distributions, and spatiality is an important and statistically accountable factor. Although there were no significant changes in density distributions and spatial dependencies when comparing data from both years, it was evidenced that population movements exhibited a more heterogeneous distribution over time in the urban space. The increase in spatial heterogeneity in 2020 was triggered not only by population movements but also by an increase in mortality rates and significant migration outflows due to the global pandemic in 2019 (Baser, 2021).

The study's usefulness lies in its requirement to examine spatiality as an effective indicator in research on density distribution in metropolitan areas such as Istanbul. Future studies could add new methods to the proposed hybrid approach to test the validity of the results. The study provides fundamental knowledge for future policy implementations regarding sustainable urban growth, as well as important insights into understanding the spatial distribution and socioeconomic dynamics of the city. Researchers can test the same methods and datasets in other metropolitan cities and compare them to our study. Additionally, they can contribute to the literature by introducing new indicators.

ETHICS: There are no ethical issues with the publication of this manuscript.

PEER-REVIEW: Externally peer-reviewed.

CONFLICT OF INTEREST: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FINANCIAL DISCLOSURE: The authors declared that this study has received no financial support.

REFERENCES

Akseki, U., Gök, B., & Deliktaş, E. (2014). City size distributions in Central Asian republics: ZIPF'S Law. Ege Academic Review, 14(2), 295–304.

Anselin, L. (1988). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. Geogr Anal, 20(1), 1–17.

Anselin, L. (1995). Local indicators of spatial association - LISA. Geogr Anal, 27(2), 93–115.

Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. Geogr Anal, 23(2), 112–131.

Arshad, S., Hu, S., & Ashraf, B. N. (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. Physica A Stat Mech Appl, 492, 75–92.

Auerbach, F. (1913). Das gesetz der bevölkerungskonzentration. Petermanns Geogr Mittheil, 59, 74–76.

Baser, O. (2021). Population density index and its use for distribution of Covid-19: A case study using Turkish data. Health Policy, 125(2), 148–154.

Bayartan, H. (2003). Geçmişten günümüze İstanbul'da nüfus. Coğrafya Derg, 11, 5–20.

Beckmann, M. J. (1999). Assignment. In Lectures on Location Theory (pp.113-120). Springer.

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. J Pol Econ, 122(3), 507–553.

Bergs, R. (2021). Spatial dependence in the rank-size distribution of cities–weak but not negligible. Plos One, 16(2), e0246796.

Black, D., & Henderson, V. (2003). Urban evolution in the USA. J Econ Geogr, 3(4), 343–372.

Brakman, S., Garretsen, H., Van Marrewijk, C., & Van Den Berg, M. (1999). The return of Zipf: Towards a further understanding of the rank-size distribution. J Reg Sci, 39(1), 183–213.

Brush, J. E. (1966). Walter Christaller. In Central Places in Southern Germany (pp. 230). Englewood Cliffs.

Casetti, E. (1972). Generating models by the expansion method: Applications to geographical research. Geogr Anal, 4(1), 81–91.

Christaller, W. (1966). Central Places in Southern Germany. Prentice-Hall.

Córdoba, J. C. (2008). A generalized Gibrat's law. Int Econ Rev, 49(4), 1463–1468.

Corominas-Murtra, B., & Solé, R. V. (2010). Universality of Zipf's law. Phys Rev E Stat Nonlin Soft Matter Phys, 82(1), 011102.

Çetin, I., & Sevüktekin, M. (2016). Türkiye'de gelişmişlik düzeyi farklılıklarının analizi. Uluslararası Ekonomik Araştırmalar Derg, 2(2), 39–61.

Deliktaş, E., Önder, A. Ö., & Karadag, M. (2013). The size distribution of cities and determinants of city growth in Turkey. Eur Plan Stud, 21(2), 251–263.

Dicken, P., & Lloyd, P. E. (1990). Location in space: Theoretical Perspectives in Economic Geography. Harper Collins Publishers.

Dittmar, J. (2009). Cities, Institutions, and Growth: The Emergence of Zipf's Law. University of California.

Dokmeci, V. (1986). Turkey: Distribution of cities and change over time. Ekistics, 53, 13–17.

Dokmeci, V., & Türk, S. S. (2001). The application of expanded rank-size model in Turkish urban settlements. In 41st Congress of the European Regional Science Association, Zagreb, Croatia.

Duran, H. E., & Cieślik, A. (2021). The distribution of city sizes in Turkey: A failure of Zipf's law due to concavity. Reg Sci Policy Pract, 13(5), 1702–1719.

Duran, H. E., & Özkan, S. P. (2015). Trade openness, urban concentration and city-size growth in Turkey. Reg Sci Inq, 7(1), 35–46.

Fischer, M. M., & Wang, J. (2011). Spatial data analysis: Models, methods and techniques. Springer Science & Business Media.

Gabaix, X. (1999). Zipf's Law and the growth of cities. Am Econ Rev, 89(2), 129–132.

Gabaix, X., & Ioannides, Y. M. (2004). The evolution of city size distributions. In Handbook of regional and urban economics (Vol. 4, pp. 2341–2378). Elsevier.

Giesen, K., & Südekum, J. (2011). Zipf's law for cities in the regions and the country. J Econ Geogr, 11(4), 667–686.

Hackmann, A., & Klarl, T. (2020). The evolution of Zipf's Law for US cities. Pap Reg Sci, 99(3), 841–852.

Ioannides, Y. M., & Overman, H. G. (2003). Zipf's law for cities: An empirical examination. Reg Sci Urban Econ, 33(2), 127–137.

Kitchin, R., & Thrift, N. (2009). International encyclopedia of human geography. Elsevier.

Knudsen, T. (2001). Zipf's law for cities and beyond: The case of Denmark. Am J Econ Sociol, 60(1), 123–146.

Kundak, S., & Dökmeci, V. (2018). A rank-size rule analysis of the city system at the country and province level in Turkey. ICONARP Int J Archit Plan, 6(1), 77–98.

LeSage, J. P. (1999). The theory and practice of spatial econometrics. Univ Toledo, 28(11), 1–39.

Lösch, A. (1954). The economics of location. Yale University Press.

Luckstead, J., & Devadoss, S. (2014). Do the world's largest cities follow Zipf's and Gibrat's laws? Econ Lett, 125(2), 182–186.

Marin, M. C. (2007). 1985 sonrası Türkiye'deki kentsel sistemin dönüşümü: Zipf Yasası'nın bir testi. Gazi Univ Muh Fak Derg, 22(1), 33–38.

Mukherji, N., & Silberman, J. (2018). Knowledge flows among US metro areas: Innovative activity, proximity, and the border effect. Rev Reg Stud, 48(2), 193–216.

Nitsch, V. (2005). Zipf zipped. J Urban Econ, 57(1), 86–100.

Parr, J. B. (1985). A note on the size distribution of cities over time. J Urban Econ, 18(2), 199–212.

Parr, J. B., & Suzuki, K. (1973). Settlement populations and the lognormal distribution. Urban Stud, 10(3), 335–352.

Rey, S. J., & Montouri, B. D. (1999). US regional income convergence: a spatial econometric perspective. Reg Stud, 33(2), 143–156.

Sahasranaman, A., & Jensen, H. J. (2020). Distribution of neighborhood size in cities. arXiv preprint arXiv:2010.06946.

Sun, X., Yuan, O., Xu, Z., Yin, Y., Liu, Q., & Wu, L. (2021). Did Zipf's Law hold for Chinese cities and why? Evidence from multi-source data. Land Use Policy, 106, 105460.

TC Resmi Gazete. (2012, November 12). Kanun no.6360: Establishment of a metropolitan municipality and the determination of its borders. TC Resmi Gazete, 28489, p. 1–3. https://www.resmigazete.gov.tr/eskiler/2012/12/20121206-1.htm

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. Econ Geogr, 46(sup1), 234–240.

TSI. (2021, April 3). Population statistics of İstanbul by neighborhood. www.tuik.com.tr.

TSI. (2024, January 21). Statistics of foreign population in İstanbul. www.tuik.com.tr.

Zeren, F. (2010). Mekansal etkileşim analizi. Istanbul Univ Ekon Ekonometri İstat e-Dergi, 12, 18–39.

Zipf, G. K. (1949). Human behaviour and the principle of least-effort. Addison-Wesley.

WHO. (2024). Death rates due to COVID-19. https://data.who.int/dashboards/covid19/cases?n=c