ARTICLE

<https://doi.org/10.1057/s41599-022-01035-5>

OPEN

Spatial stratification and socio-spatial inequalities: the case of Seoul and Busan in South Korea

Seungwoo Han ¹✉

This study approaches the spatial stratification phenomenon through a data-based social stratification approach. In addition, by applying a dissimilarity-based clustering algorithm, this study analyzes how regions cluster as well as their disparities, thereby analyzing socio-spatial inequalities. Ultimately, through map visualization, this study seeks to visually identify spatial forms of social inequality and gain insight into the social structure for policy implications. The results determine how the regions are socioeconomically structured and identify the social inequalities between the spaces.

¹Rutgers University, Newark, NJ, USA. ✉email: seungwoo.han@rutgers.edu

Introduction

As social inequality worsens worldwide, its manifestation in complex urban environments has become a key issue in policy research. Many studies on urban inequality have attempted to measure inequality by combining the level of the economy, income, education, public service, and life expectancy (Alkire et al., 2011; Lee and Rodri'guez-Pose, 2013; Panori and Psycharis, 2017; Lelo et al., 2019). Other studies use non-material factors such as the perception of quality and happiness of life (Senlier et al., 2008; Ballas, 2013; Okulicz-Kozaryn, 2013). In the wider context of overall social inequality in regard to space, however, we need to develop a better understanding of the mechanisms that shape socio-spatial inequality. In order to analyze the spatial patterns of social inequality, this study focuses on the opportunities and benefits coming from space and measures spatial stratification by analyzing the multifaceted factors that create disparities among spaces.

The objectives of this study are twofold. The first is a methodological discussion of perspectives, approaches, and data to measure spatial stratification by applying data-driven methods. The other is the application of this approach to understanding socio-spatial inequalities due to spatial stratification in South Korea. This study covers Seoul Special City (hereinafter referred to as Seoul), the capital city of South Korea, and Busan Metropolitan City (hereinafter referred to as Busan), the second-largest city in South Korea. The units of analysis are the district (*gu*) and county (*gun*) to which the two cities belong.

The inequality referred to in this study is social inequality. Social inequality refers to a state in which factors affecting human activities across various fields, such as opportunities, resources, and power, are unfairly distributed (Sen, 1992). Socio-spatial inequality, then, refers to a state in which significant disparities are created because they are not evenly distributed across different spaces, which means that social inequalities are manifested in spatial patterns. It proposes that socio-spatial inequalities can be identified by measuring spatial stratification.

The approach to spatial stratification in this study is based on understanding social stratification via data-driven methods. As a research method for clustering regions and analyzing disparities, this study proposes the *K*-means ++ clustering algorithm, which is a dissimilarity-based (distance-based) clustering method from a problem-centric perspective of spatial stratification and socio-spatial inequalities.

This study presents interpretable clustering results through a combination of a clustering algorithm and map visualization for policy implications. In the study of socio-spatial inequalities in urban spaces, the approaches using map visualization enable the spatial analysis of urban inequalities to visually identify spatial forms of inequalities, gain deeper insight into social structure and the processes that generate inequalities (de la Espriella, 2009; Soja, 2010; Siqueira-Gay et al., 2019; Lelo et al., 2019; Sohn and Oh, 2019; McLachlan and Norman, 2020; Shi and Dorling, 2020). The results of this analysis can be used as a foundation in policy discussions related to urban and regional inequalities and this study seeks to find implications through this approach.

In a study on data-based social stratification, it is crucial to choose which indicators meet the research objective. This study proposes that data reflecting the multifaceted characteristics of spaces have a certain pattern to measure spatial stratification. This study uses data, which reflect country-specific characteristics, provided by 15 public institutions in South Korea for its analysis. In addition, this study applies data transformations that can effectively maximize similarity and dissimilarity to optimally cluster regions based on the dissimilarity-based clustering method. As tools for analysis, this study used Python 3.7 and Scikit-learn 0.22.2.

Socio-spatial Inequality

The spatial organization of urban inequality. Many studies heavily rely on income data to identify spatial inequality even though it is widely acknowledged that inequality is a multifaceted phenomenon affecting human activities across various fields (Sen, 1992). Besides income, studies to look at the socioeconomic structure of a specific space have focused on occupation, housing, and education (Jung et al., 2014; Kernan and Bruce, 1972; Henning and Liao, 2013; Sohn and Oh, 2019). They certainly help to understand the socioeconomic structure of space, but they do not show the process by which inequality is (re)produced. Urban inequality is multidimensional and highly complex. Multidimensional analysis of space provides a different perspective on its socioeconomic structure (Hacker et al., 2013; Lelo et al., 2019; Lin et al., 2015; Nijman and Wei, 2020; Spector, 1982; Zambon et al., 2017).

The central theme of this study of socio-spatial inequality is the spatial organization of urban inequality. This study argues that the spatial arrangement of economic and service facilities and classes helps us understand how space is structured socio-economically. First, socio-spatial inequality is derived from the spatial arrangement of economic and service facilities related to the lives of residents in a city. George (1973) attempted to determine why poverty and inequality have not ended despite the progress of society. He found the answer in land. According to his claims, progress is beneficial to humanity and also increases the value of the land, so the amount of rent that the landlord can demand those who need to use the land also increases. The landlord monopolizes the fruit of growth because the price of land and the rent charged for using it increase faster than the increasing wealth to pay that rent. In other words, inequality intensifies as landowners become increasingly able to monopolize the surplus arising from economic growth.

When we speak of land, we refer not to soil and stone but to a specific location. Due to the increase in population and settlement activities following the progress of society, the scarcity of land becomes greater, so the value of land is determined by the location. To put this in terms of modern society, the primary role of land for individuals in contemporary society is the role of housing owned by individuals. Housing provides social benefits and opportunities beyond the purpose of residence. Facilities occupy discrete locations, and the friction of distance means that some people in certain places will find it easier than others to obtain opportunities, benefits, and various sources of need satisfaction (Smith, 1984). Individual behavior is affected by spatial structure (Horton and Reynolds, 1971). The residential location carries with it not only a particular quality of living environment but also a set of advantages and disadvantages arising from accessibility to sources of benefits and opportunities (Su et al., 2019). Disadvantages arising from lack of accessibility to sources of benefits and opportunities might affect economic performance, which might reproduce inequality (Rawls, 1971; Roemer, 1998; Lamont and Fourier, 1992; Thobcke and Charumilind, 2002; Mustard and Ostendorf, 2005; van Kempen, 2005; McDowell et al., 2006). In other words, the key to socio-spatial inequality in this study is the socio-spatial inequality of opportunities and benefits.

Second, the socio-spatial inequality of opportunities and benefits stems from spatial exclusivity. This exclusivity is closely related to the traits of the positional good discussed by Veblen (1994) in that the location of a particular region represents the class of the individuals residing there. People pay more for houses in locations that are coveted in terms of social status, and they are willing to bear it even if they account for more of their assets. As a result, higher barriers to entry are built up, and those who do not

belong are excluded from the benefits and opportunities of the region.

In social science studies, class is an essential factor in explaining society's various dynamics and phenomena regarding inequality, from classical discussions about the class such as Marx (1977) to Piketty (2014, 2020), who discusses class in terms of the present time. The distribution of class shows the socioeconomic structure of society (Reich, 1991; Atkinson, 2006, 2008; Sohn and Oh, 2019). Class and social strata exist in any form, regardless of age and place, and inevitably, the accompanying inequality is a byproduct of their dynamics (Pekkanen et al., 1995; Chan and Goldthorpe, 2007; Kingstone, 2000; Wright, 2005; Grusky, 2014; Piketty, 2020). If inequality exists in any form and social relations between people are perceived as non-horizontal by any standard, class and social strata can be useful tools for analyzing socio-spatial inequality.

The Korean context. In South Korea, phrases that represent specific spaces, such as the metropolitan area versus rural provinces, in-Seoul versus out-of-Seoul, and Gangnam¹ versus Gangbuk reflect individuals' identity, social status, and economic class (Kang, 1991; Park and Jang, 2020; Yang, 2018). Phrases that define regions in specific ways mean that spatially, social classes are rigidly separated and the opportunities available to individuals vary depending on where they live. Whether an individual lives in a metropolitan area, in a rural area, or in Gangnam within Seoul, affects one's life in South Korean society in many ways. Inequality can be structurally reproduced and social mobility becomes rigid if a certain group of people living in a certain area monopolizes opportunities, or if some people are spatially excluded from opportunities provided by society (Soja, 2010).

Previous literature on the regional inequality of South Korea mainly dealt with the economic gap between metropolitan and rural areas (Kim et al., 2003; Kim and Jeong, 2003; Noh, 2006; Oh, 2017). On the other hand, the core of structural inequality in South Korea can be captured through the analysis of Seoul and Busan in that multifaceted inequality factors are concentrated in the space of these two representative cities in South Korea. As cities increase in size, diversity also increases and reveals the overall social structure of society (Shevky and Bell, 1955; Duranton and Puga, 2000). As of 2020, Seoul's population is about 9.8 million, accounting for about 18.8% of the total population of South Korea (about 51.84 million), and Busan's population is ~3.5 million, accounting for about 6.6% of the total population of South Korea;² taken together, the two constitute more than ¼ of South Korea's total population. Accordingly, urban inequality in Seoul and Busan is not limited to the urban space but, rather, can show the overall structure of regional inequality in South Korean society.

According to the Global Power City Index (GPCI), Seoul is ranked 8th³ and according to the Global Cities Index (GCI), Seoul is ranked 17th among global cities in 2021.⁴ In terms of container traffic per annum, Busan is ranked 6th in the world and is considered one of the key cities for port logistics in 2021.⁵ Therefore, socio-spatial inequalities within the two cities, which play a key role socially and economically and are closely linked to the global economy, can be understood as a form of inequality in the global city.

Research design

Data-based social stratification approach. The approach to spatial stratification in this study is based on measuring social stratification through data-driven methods. Data reflect human behaviors and interactions, such as how people communicate, how they form relationships, and how conflicts arise in society

(Monroe et al., 2015). People live by building complex dynamics of life inside and outside each spatial unit. The spatial units of various lives contain each way of life and relationship. The spatial units are created by human beings as the main subject through social relations, but at the same time, society also creates spatial units through institutional or relational networks. As such, the dynamics of confrontation and rejection as well as connection and bonding are laid between each unit space. If the social stratification phenomenon reflects people's social relations, the spatial stratification phenomenon reveals these social relations as spatial divisions. Accordingly, this study proposes that the multidimensional data reflecting social relationships have a certain pattern by which to measure spatial stratification. Table A1 of Appendix A summarizes the geographical scope, methods, indicators, and findings of past publications of applied clustering.

Dissimilarity-based clustering methods. Previous studies have generally applied one or two certain clustering algorithms for analysis (see Table A1 of Appendix A). However, there is no foundation in statistical theory or clear criteria for which clustering algorithm is preferable (Venables and Ripley, 2002; Ahlquist and Breunig, 2012; Hennig, 2015). There are a number of clustering algorithms, and, often, different methods produce different outcomes without sound reasons for choosing a particular method over another. Therefore, in selecting a clustering algorithm, it is difficult to clearly explain which is preferable and how many clusters are ideal. In a number of studies applying clustering algorithms, the reason for selecting a specific clustering algorithm is not clearly presented or discussed (Ahlquist and Breunig, 2012; von Luxburg et al., 2012).

This study does not aim to compare each result by applying various clustering algorithms. To determine which clustering method is preferred and suitable for clustering, the current study takes an approach in which the researcher determines the clustering algorithm to be applied in accordance with the objective and context of the research as well as the characteristics of the data (von Luxburg et al., 2012; Henning and Liao, 2013; Henning, 2015). Therefore, the current study is based on the data-driven approach rather than the model-driven approach.

Each region within the urban space is unique, so the regional characteristics of each region are different (Harvey, 1989). At the same time, however, certain regions share unique features based on specific values. Current study proposes a dissimilarity-based clustering method by focusing on this similarity and dissimilarity as reflected in data in order to measure spatial stratification. As discussed in the previous section, this study proposes that multifaceted data have a certain pattern that can be utilized to measure spatial stratification. According to this data-based social stratification approach, structural patterns can be elucidated. This study seeks to uncover them based on the similarity and dissimilarity among observations. Accordingly, this study applies the *K*-means++ clustering algorithm, an approach to clustering based on Euclidean distance (see Appendix B).

In addition to *K*-means++, there are clustering algorithms of various approaches, such as hierarchical clustering and density-based spatial clustering with noise (DBSCAN). Hierarchical clustering has the advantage that it can determine the number of clusters by searching all potential clusters through a hierarchical tree structure (Murphy, 2012; Johnstone et al., 2019; Wu et al., 2020). In hierarchical clustering, clusters have a tree-like structure or a parent-child relationship. Here, the two most similar clusters are joined together, and all of the clusters are continuously combined until they form a single cluster. DBSCAN is a density-based clustering method that is a non-parametric approach suitable for applications where clusters cannot be well described

as distinct groups of low within-cluster dissimilarity, as, for instance, in spatial data, where clusters of points in the space may form along natural and artificial structures, such as rivers, valleys, buildings, etc. (Grubestic et al., 2014; Henning, 2015; Johnstone et al., 2019; Wu et al., 2020). The objective of this study is not to connect objects hierarchically to multiple clusters but to directly optimize certain characteristics and categorize each object into exactly one cluster. In addition, this study's data do not require a density-based method because geographic characteristics are not included.

K-means++. K-means++ is one of the clustering algorithms developed from K-means, and the principle of clustering is the same except for the initialization of the cluster center. K-means is a clustering technique that selects a cluster center called a centroid and then selects the data points closest to it (Arthur and Vassilvitskii, 2007; Hastie et al., 2009; Murphy, 2012) (see Appendix C).

The main disadvantage of K-means is that the initial locations of centroids are arbitrarily selected. This initial arbitrary selection of centroids often fails to form optimal clusters. K-means++ is the clustering algorithm proposed to address this drawback of K-means (Arthur and Vassilvitskii, 2007; Bonaccorso, 2018). It specifies a procedure to initialize centroids before moving forward with the standard K-means clustering algorithm.

K-means performs the clustering process by initially arranging random centroids. In contrast, K-means++ selects one of the data points as the first centroid, rather than beginning with K points in arbitrary spaces. It then selects the next centroid from the data points such that the probability of choosing a point as a centroid is directly proportional to its distance from the nearest, previously chosen centroid (Arthur and Vassilvitskii, 2007). Simply put, a data point placed as far as possible from the already designated centroid is designated as the next centroid. This process is repeated until K centroids have been sampled. In other words, initial centroids are placed more strategically rather than randomly selected in the centroid selection. Except for this initial procedure, the rest of the clustering process is the same as K-means. The approach of K-means++ to initial centroid selection can cluster objects more optimally and improve the algorithm's convergence speed.

In K-means++ clustering, the number of clusters K must be specified before clustering. That is, what must be decided here is how many clusters K are optimal. Silhouette analysis can be used to evaluate the separation distance between the resulting clusters (Kaufman and Rousseeuw, 1990; Bonaccorso, 2018). Efficiently clustered means that the distances between different clusters are sufficiently far apart, and data points in the same clusters are close. The silhouette plot displays a measure of how close each data point in one cluster is to data points in the neighboring clusters and thus provides a way to assess parameters such as the number of clusters visually (see Appendix D).

Data selection. A critical question for the data-based social stratification approach is what indicators to choose. When collecting data, it is necessary to have a sufficient understanding of the society concerned, and data should be available and reliable. For the data set that is analyzed here, the focus is on economic and service facilities and socioeconomic class. Data related to the spatial arrangement of economic and service facilities include data representing the sectors of transportation, culture, safety, medical treatment, education, and economy. Class includes data related to an individual's socioeconomic level, such as educational background, occupation, income, and wealth (Hollingshead, 1975; Levy and Michel, 1991; Sohn and Oh, 2019).

In modern society, public transportation plays a role in distributing opportunities to people through mobility (Social Exclusion Unit, 2003; Lucas, 2012; Chen et al., 2018; Pizzol et al., 2021). Among the various means of public transport, in South Korea, the subway is considered the most essential for urban transportation (Im and Hong, 2017). According to the Seoul Metropolitan Government, the average number of subway passengers per day is over 5 million, surpassing other modes of public transportation.⁶ In South Korea, the area around a subway station is called a "subway station influence area", and considering the fact that commercial areas, businesses, and public institutions are located and various social and economic activities take place near subway stations, subway stations are more important than being merely a means of transportation in various ways. In addition, considering the direct and indirect effects of the transportation infrastructure on the region and the parking problems in Korean metropolitan areas, public investment in roads and public parking spaces are also essential factors for residents (Talley, 1996; Yi et al., 2012; Ahn et al., 2014).

Cultural facilities such as public libraries, museums, and art galleries form cultural capital and are essential elements affecting the quality of life, vitality, and performance of individuals (Andersen and Hansen, 2012). In South Korea, cultural facilities have essential meanings in terms of quality of life, regional vitality, and the competitiveness of residents (Kim, 2007; Park et al., 2015). There has been continuous discussion regarding the disparities in accessibility to such facilities. In addition, access to a movie theater is one of the key factors in increasing the overall level of cultural activities in a region. In South Korea, the multiplex cinemas, which account for more than 90% of total cinemas,⁷ provide the concept of a comprehensive leisure facility that can be enjoyed not only for movies but also for other leisure activities (Kang, 2016).

In the case of medical care, in South Korean society, there are health inequalities within and between regions (Choi et al., 2011; Hong and Ahn, 2011). In particular, tertiary hospitals occupy an important position such that the unique term "tertiary hospital influence area" was necessitated (Kang, 2014). There has been continuous social debate on patients' inclination toward the top five tertiary hospitals located in Seoul. Considering the high medical service level of tertiary hospitals, residents can enjoy high levels of benefits (Yang et al., 2020). Safety needs are important factors for residents' lives in modern society (Cox and Cox, 1996). Regarding the safety of residents in South Korea, there has been constant discussion that the utility level for people's safety differs depending on accessibility to CCTV, police stations, and firehouses (Kim, 2014).

The distribution of educational opportunities as well as access to them has become an important issue in relation to educational and social equality (Coleman, 1990; Talen, 2001; Zhang and Kanbur, 2005). The concept of equality in educational opportunities includes the right for students to receive the benefits of a common curriculum regardless of their social background as well as the right to equal education in the community (Coleman, 1990). Considering the social phenomena that education is projected as a desire to increase social status in South Korean society as well as of parents' enthusiasm for their children's education, the meaning of education is highly significant (Seth, 2002; Lee, 2005; Kang, 2008).

In South Korea, disparities in educational services among regions are discussed as a serious social problem (Son, 2004; Choi, 2004; Byun and Kim, 2010; Byun et al., 2012). In particular, the disparities in the enrollment rates of elite high schools, such as specialized high schools and autonomous private high schools, which are advantageous for entering major universities, between regions are significant. In addition, according to the National

Table 1 List of variables.

	Category	Sub-category	Variable
1	Economic and service facility	Transportation	Subway station
2			Public parking spaces
3			Road extension (m)
4			Road extent (m ²)
5		Culture	Cultural facility (Public library, Museum, Art gallery, Art center, Local culture center, etc.)
6			Theater
7		Medical treatment	Tertiary hospital
8		Safety	CCTV
9			Police station (Police substation, Community security center)
10			Firehouse
11		Education	Enrollment rates of elite high schools
12			Private educational institute
13		Economy	Large-scale stores (Super Super Market (SSM), Department store, Shopping center, Multi-shopping complex, etc.)
14	Class		Gross wage and salary (based on the location of withholding agent) (unit: million KRW)
15		Education	High educational background
16		Occupation	High professional skill
17		Income	High income
18		Wealth	The average price of a condominium (unit: 1000 KRW)

See Tables E1–E4 in Appendix E for more detailed data and descriptive statistics.

Fig. 1 Transformation $\log(x + 1)$. **Fig. 2** Transformation $\log(x + c)$.

Statistical Office's announcement in 2019, 82.5% of elementary school students, 69.6% of middle school students, and 58.5% of high school students were receiving private education.⁸ In this respect, the proportion of each district in the city's total elite high school enrollment and the number of private educational institutes are included for the analysis.

Local shops are closely related to residents' demographic characteristics (Meltzer and Schuetz, 2011). In Korean society, large-scale stores, such as super super market (SSM), department stores, shopping centers, multi-shopping complexes, etc., are factors that affect the residents' quality of life (Kim and Park, 2017). In the case of the regional economy, the district's gross wage and salary based on the withholding agent's location show the region's overall level of economic activity and job opportunities (Chapple, 2007). The high gross wage and salary of a district imply its economic competitiveness.

This study's data representing class include the level of residents' education, professional skills, income, and wealth. Educational background, occupation, income, and wealth are representative factors of socioeconomic class (Hollingshead, 1975; Kim et al., 2003; Sohn and Oh, 2019). In terms of educational

background, university (including vocational college) graduation or above is classified as high, and high school graduation or below is classified as low. Based on the Korean Standard Classification of Occupations, professional or higher is classified as high, and others are classified as low for professional skill level. In regard to income, the fourth quartile is classified as high, and the first quartile is classified as low. The ratio of the working population of the upper tier to the lower tier in each data is measured based on national census data (KEIS, 2019).

This study uses the price of a condominium (called an apartment in South Korea) as data representing an individual's wealth. According to the Korea Housing Survey of the Ministry of Land, Infrastructure and Transport, in Seoul, as of 2018, about 42% of households live in condominiums. In Busan, about 53.6% of households live in condominiums.⁹ In South Korean society, the price of condominiums is heavily influenced by the region in which they are located and the surrounding living environment, and this is one of the main factors that characterize the wealth and socioeconomic status of an individual (Zhang, 1998; Lee et al., 2002; Choi, 2006; Lee, 2009; Jang and Kang, 2015; Sohn and Oh, 2019). The variables are shown in Table 1.

Fig. 3 Socio-spatial map of Seoul 1.

Data transformations. This study applies data transformations that can effectively maximize similarity and dissimilarity in order for regions to optimally cluster by applying a dissimilarity-based clustering algorithm. From a data-intuitive perspective, it may be meaningful to find a pattern from data without transformations, but this study considers that it makes more sense to cluster by ratios through log transformations rather than relying on absolute differences in variable values in that, in terms of social stratification, the interpretive difference between social groups depends on ratios rather than absolute values (Henning and Liao, 2013). In this study, therefore, the log transformations are applied to all variables except for the ratio variables.

Since there are 0s in the data, the transformation $\log(x+c)$ is appropriate. The strategic consideration in selecting c is that, rather than adding 1 to x uniformly, adding each corresponding c considering the minimum and maximum values of each variable enables more efficient clustering. For example, in the number of movie theaters in Seoul, the minimum value is 0, and the maximum value is 9. In contrast, for gross wage and salary, the minimum value is 972,996 (unit: 1 million KRW), and the maximum value is 34,245,070 (unit: 1 million KRW). Accordingly, it is logically appropriate to select c to be applied to the movie theaters variable and c to be applied to the total gross wage and salary variable differently.

The selection of c considering the values of variables is subjective, and this study takes a method of adding a multiple of 10, which is one digit greater than the maximum value. This makes the distance between small values effective while leaving the effective distance between high values less affected. Figures 1 and 2 are an example of clustering according to the difference in c values in the transformation $\log(x+c)$. This shows the difference between Fig. 1 the case of applying 1 to c and Fig. 2 this study's approach when clustering with the average price of a condominium and the number of private institutes. The approach of this study makes clustering more efficient.

Clustering results and analysis

Before examining the clustering results, we can briefly analyze the socio-spatial maps of Seoul and Busan in Figs. 3–6, which deliver multifaceted aspects of the socio-spatial structures in an intuitive visual manner. In Figs. 3 and 4, we can visually confirm that the elements constituting transportation, culture, safety, education, and economy are concentrated in the south of the Han River, Seoul. Looking at the class factors, it can be seen that the corresponding factors are very high in the south of Seoul compared to other regions.

In terms of accessibility to the facilities for residents, the facilities providing opportunities and benefits are concentrated in

Fig. 4 Socio-spatial map of Seoul 2.

the southern area of Seoul, and highly educated, professional, high-income, and wealthy social classes reside in the area. In the following, we can look at Fig. 5 which briefly shows the socio-spatial structure of Busan.

In the case of Busan, compared to Seoul, the concentration of elements constituting transportation, culture, safety, education, and economy in a specific region is relatively weak. However, many facilities are still concentrated in the southeast region (East Busan). Gross wage and salary are higher in the west. This may be because Busan's port facilities and related businesses are located in the west. In terms of social class, more of highly educated, professional, high-income, and wealthy social classes reside in the southeast region compared to other regions.

From the above maps, we can see the socio-spatial structures of the two cities. In the following, we can further understand their socio-spatial inequalities by analyzing the clustering results. The results of the silhouette analysis of Seoul and Busan are shown in Figs. 6 and 7, respectively. In Seoul, when divided into two clusters ($K=2$), the silhouette score is the highest (0.59). On the other hand, as K increases to three (0.511), four (0.445), five (0.437), and six (0.415), the silhouette score decreases gradually. In the case of Busan, when $K=4$, it has the highest silhouette score (0.364). In the case of $K=2$ and $K=3$, clustering is not efficient because there are clusters with negative values, and as K increases to five (0.3) and six (0.255), the silhouette score

decreases gradually. Therefore, $K=4$ seems to be the most appropriate. First, Fig. 8 shows a map visualization of the clustering result for Seoul ($K=2$).

We can see the spatial shape of the result clustered into two clusters in Fig. 8. The districts included in each of the two clusters in Seoul are shown in Table 2. In the case of Seoul, 22 districts form Cluster 0, and three districts form Cluster 1. That is, Gangnam, Seocho, and Songpa districts form one cluster, and the rest of the districts form the other cluster.

The disparities in the mean values between the two clusters can be clearly distinguished. In all respects, Cluster 1 has overwhelming advantages. Considering subway stations, the average number in Cluster 0 is ~12.7, and the average number in Cluster 1 is 26.3. In the case of public parking spaces, Cluster 1 has about twice as many spaces on average. There is relatively little difference between the two clusters in terms of road extension, but in the case of the road extent, the difference is ~1.8 times. In the case of cultural facilities, there are about 12 cultural facilities on average in Cluster 0, but in Cluster 1, there are about 21 cultural facilities on average. In the case of theaters, there are about 2.7 theaters on average in Cluster 0, but in Cluster 1, there are about 3.5 theaters on average.

In the case of tertiary hospitals, each district of Cluster 1 has at least one, but in Cluster 0, the average number of tertiary hospitals is less than zero. For safety, Cluster 1 has at least 1.3 times

Fig. 5 Socio-spatial map of Busan.

more CCTVs, police stations, and firehouses on average. Regarding education, the gap between the two regions is considerable. There is a sizable gap between the two clusters in the enrollment rates of elite high schools and the number of private educational institutes. In the economy, the average number of

large-scale stores in Cluster 1 is about 1.8 times higher, and the average gross wage and salary of Cluster 1 are over 4.6 times higher. In terms class, Cluster 1 significantly exceeds Cluster 0 in all areas of education, professional skill, and income. The average price of a condominium in Cluster 1 is about three times higher

Fig. 6 Silhouette scores ($K = 2-6$), Seoul.

Fig. 7 Silhouette scores ($K = 2-6$), Busan.

Fig. 8 Map visualization of clustering ($K = 2$), Seoul.

than that in Cluster 0. Looking at Fig. 8 and Table 3 together, we can see the spatial shape of the clusters and the disparities between them.

In Busan's case, looking at the map visualization in Fig. 9 and Table 4 of the clustering result, we can see how the regions form clusters and take a spatial shape. When $K=4$, the regions belonging to each cluster are listed in Table 4. From the following results, we can see that Haeundae district forms one cluster, six districts adjacent to the left of Haeundae form Cluster 1, and seven districts located on the left form Cluster 0. Gangseo district and Gijang County, located at both ends of Busan, form Cluster 3.

In Busan, the disparities among clusters are not relatively large compared to in Seoul. However, Cluster 2 (Haeundae district) is superior in most sectors, except for tertiary hospitals. Cluster 2 shows that the number of subway stations is two to three times

higher than in other clusters, and the number of public parking spaces is greater. Regarding the number of cultural facilities, it is about two times higher than that of other clusters, and the number of movie theaters in Cluster 2 is 6–7 times higher than that of others. The number of police stations is similar to that of Cluster 1, but it is about two times higher than that of others. In the case of firehouses, the number is more than twice that of Cluster 1 (Table 5).

In the case of enrollment rates of elite high schools and the number of private educational institutes, Cluster 2 greatly exceeds other clusters. The number of large-scale stores in Cluster 1 is at least two to four times higher than the other clusters. Considering gross wage and salary, the gap with Cluster 3 is not significant, but it is about 1.7 times higher than that of Cluster 0. In terms of class, looking at the gaps in education, professional skill, and income, these gaps are not significant, but they clearly exceed

Fig. 9 Map visualization of clustering (K = 4), Busan.

Table 2 Districts by cluster (Seoul).		
Cluster	Districts (Gus)	Numbers
Cluster 0	Dobong, Dongdaemun, Dongjak, Eunpyeong, Gnagbuk, Gangdong, Gangseo, Geumcheon, Guro, Gwanak, Gwangjin, Jongno, Jung, Jungnang, Mapo, Nowon, Seodaemun, Seongbuk, Seongdong, Yangcheon, Yeongdeungpo, Yongsan	22
Cluster 1	Gangnam, Seocho, Songpa	3

Table 3 Mean value of variable by cluster (Seoul).				
	Sub-category	Variable	Cluster 0	Cluster 1
1	Transportation	Subway station	12.72	26.32
2		Public parking spaces	6534.79	13717.45
3		Road extension (m)	306,319.4	397,962.4
4		Road extent (m ²)	3,086,839	5,408,973
5	Culture	Cultural facility	12.48	21.21
6		Theater	2.74	3.58
7	Medical treatment	Tertiary hospital	0.31	1.29
8	Safety	CCTV	1954.38	2863.44
9		Police station	16.75	22.26
10	Education	Firehouse	4.41	6.00
11		Enrollment rate of elite high school	3.17	9.77
12	Economy	Private educational institute	311.94	1146.15
13		Large-scale store	15.55	27.03
14		Gross wage and salary (unit: million KRW)	4,255,825	19,541,940
15	Educational background	High educational background	52.17	70.32
16	Occupation	High professional skill	28.06	42.77
17	Income	High income	25.99	43.49
18	Wealth	Average price of a condominium (unit: 1000 KRW)	554,130	1,303,537

See Fig. E1 in Appendix E for graphs.

other regions in all these areas. The price of a condominium in Cluster 2 is about twice as high as in Cluster 0.

In summation, through the analysis of the clustering results, we can identify the spatial patterns of social inequality. Certain regions, densely populated by socioeconomically upper-class people, offer residents higher levels of benefits and opportunities than other regions. In conclusion, through these findings, this study is able to determine how the regions are socioeconomically structured spatially and to identify the social inequalities between the spaces.

Conclusions and implications

This study has several main findings, based on the methodological discussion that addresses a series of views on the perspectives, approaches, and data. In Seoul, the highest average silhouette score is calculated when divided into two clusters, and in Busan, the clustering is most optimal when divided into four clusters. Seoul’s Cluster 1 has advantages over other clusters in all sectors of economic and service facility, and class. As a result, this group’s residents can enjoy higher levels of services of public transportation, safety, medical treatment, culture, education, and

economic opportunities and benefits compared to other regions. In the case of Busan, Cluster 2 has advantages over other clusters in most sectors of economic and service facility, and class. Compared to Seoul, the degree of disparity among clusters is relatively small. Still, there are evident disparities in the benefits and opportunities between them. Obviously, certain regions, densely populated by socioeconomically upper-class people, offer residents higher levels of benefits and opportunities than others.

Before stressing the broader implications, it is necessary to be clear about the theoretical and empirical limitations of this analysis. The proposed causal explanation linking location, benefits and opportunities, class, and socio-spatial inequality is tentative and begs further exploration. Empirically, the findings of this study can only be suggestive. In terms of the data-driven approach, the current study acknowledges some degree of arbitrariness in the selection of data. Although the current study utilized available data reflecting multidimensional characteristics of inequality, there were missing parts that this study could not address because of the unavailability of data. If time-series data were available, we could look at the changes in socio-spatial structures. However, time-series data were not available either. Future research would greatly benefit from more extensive and reliable time-series data.

Methodologically, this study applied *K*-means++ in the context of the study because based on the data-driven approach it was determined that there was less need to compare and analyze the results by applying various clustering algorithms. In a follow-up study, nevertheless, it is necessary to compare various clusters to which various clustering algorithms are applied for more comprehensive interpretations.

Besides, as discussed at the beginning of this study, socio-spatial inequality refers to the state in which opportunities, resources, and power are not distributed evenly across different spaces. This study captures socio-spatial inequalities in

opportunities and resources but cannot capture political inequality from a spatial aspect. In addition, although this study makes it possible to identify the social inequalities between spaces in Seoul and Busan, it is not for the whole country. The social inequalities between Seoul and other regions may be incomparably larger than those within Seoul (Kang, 1991; Kim and Jeong, 2003; Yea, 2000). These gaps are expected to be filled through future studies.

Nonetheless, this first attempt to uncover socio-spatial inequalities in South Korea based on data-driven methods is provocative. There are many different perspectives and positions on the analysis of inequality. Previous literature on regional inequality in South Korean society has generally focused on income inequality between provinces or metropolitan cities and provinces. On the other hand, the current study analyzed how the disparities in opportunities and classes stratify urban spaces.

We need to think about what the clustering results imply. The results of this study adequately reflect the reality of South Korean society. A Korean proverb states, “The young of a human should be sent to Seoul.” This is because people can find more opportunities and benefits in big cities such as Seoul. After belonging to the space of Seoul, people want to live in a certain area, Gangnam. Similarly, in Busan, people want to move from West Busan to East Busan, where Haeundae district is located. This social phenomenon is due to the apparent existence of socio-spatial inequalities, and many people in South Korea desire to belong to the group of people living in Seoul Gangnam and Busan Haeundae. On the other hand, these regions are a space of jealousy and frustration and are often indicated as a symbol of inequality in South Korean society due to socioeconomic polarization. In other words, these regions are a space of love and an object of desire on the one hand and space of envy and frustration on the other.

According to Soja's (2010) conceptualization of spatial justice, if the geographic space formed by the social process is not socially just (it is not fair to all), the space formed in this way affects the society and lives of individuals in unjust ways. That is if spatial classes are formed in the historical moment and social context, the majority of human activities, except for a certain group of people, are spatially excluded from public services and investments. The results of this study, which targets two representative cities in South Korea, can be said to be an example to partly explain. It is worth noting that, in particular, the gap in the enrollment rates of elite high schools is significant between Gangnam and the rest of the region in Seoul, and between Haeundae and the rest of the region in Busan. This is a result that

Table 4 Districts (and county) by cluster (Busan).

Cluster	Districts (Gus)	Numbers
Cluster 0	Buk, Sasang, Saha, Seo, Jung, Dong, Yeongdo	7
Cluster 1	Geumjeong, Dongnae, Yeonje, Busanjin, Nam, Suyeong	6
Cluster 2	Haeundae	1
Cluster 3	Gangseo, Gijang	2

Table 5 Mean value of variable by cluster (Busan).

Sub-category	Variable	Cluster 0	Cluster 1	Cluster 2	Cluster 3
1	Transportation				
2					
3	Culture				
4					
5	Medical treatment				
6	Safety				
7					
8	Education				
9					
10	Economy				
11					
12	Educational background				
13	Occupation				
14	Income				
15	Wealth				

See Fig. E2 in Appendix E for graphs.

makes it possible to see that social classes are being reproduced through education in South Korean society.

This research has obvious implications at the local public policy level. Discussing and solving social problems arising from social inequality begins with a clear perception of reality. In this regard, through the findings of this study, we are able to identify which social inequality factors are interspersed between spaces and determine the spatial shape. Although it may not be possible to address the multifaceted inequalities presented in this study easily, geographical expansion of opportunities can be one of the solutions. This can be possible not from a non-spatial policy perspective, but by expanding the geography of opportunities to improve access to opportunities in specific living areas. This study's policy implications include the necessity of introducing measures to reduce the gap in opportunities and benefits between regions. Social inequality is structurally reproduced if a certain social class living in a certain area monopolizes opportunities and benefits. Therefore, how to distribute these opportunities and benefits more fairly is at the heart of policy. However, the real challenge is how to decentralize economic and service facilities that have a strong centripetal tendency. The realistic plan is to develop Seoul and Busan into a multi-centric cities. Seoul, where about 9.8 million people live, should not be a simple structure that can be divided into Gangnam and the rest (Haeundae and the rest in Busan), but a multi-centric structure in which various small and medium-sized cities are connected by education, culture, transportation, and industry. These factors should not be concentrated in one place but should be spread across regions. In other words, the current mono-centric city must develop into a multi-centric city. Since social investment in the supply of these services and facilities is difficult in the short term, it should be planned and developed from a long-term perspective. Thus, follow-up studies should investigate them further.

Data availability

All data analyzed are contained in Appendix E included in the supplementary information.

Received: 3 June 2021; Accepted: 6 January 2022;

Published online: 19 January 2022

Notes

- There is no standard that clearly defines Gangnam, but here, Gangnam is a kind of proper noun expression referring to three districts, Gangnam-gu, Songpa-gu, and Seocho-gu, called the Gangnam 3 gu. Gangnam is defined as an area that symbolizes social classes, political behavior, wealth, consumption behavior, condominium prices, educational conditions, and public services that are different from other regions (Park and Jang, 2020; Yang, 2018).
- Ministry of the Interior and Safety, Republic of Korea (<http://27.101.213.4/>).
- The Mori Memorial Foundation (<https://mori-m-foundation.or.jp/english/ius2/gpci2/index.shtml>).
- Kearney (<https://www. Kearney.com/global-cities/2021>).
- Marine Insight (<https://www.marineinsight.com/know-more/container-ports-and-port-operators/>).
- Seoul Metropolitan Government (<http://news.seoul.go.kr/traffic/archives/31616>).
- Korean Film Council (<https://www.kofic.or.kr/kofic/business/board/selectBoardDetail.do?boardNumber=2andboardSeqNumber=48560>).
- Statistics Korea (http://kostat.go.kr/portal/korea/kor_nw/1/6/1/index.board?bmode=readandaSeq=374490andpageNo=androwNum=10andamSeq=andsTarget=andsTxt=).
- Ministry of Land, Infrastructure and Transport (<http://stat.molit.go.kr/portal/cate/statFileView.do?hRsId=327&hFormId=>).

References

Ahlquist JS, Breunig C (2012) Model-based clustering and typologies in the social sciences. *Political Anal* 20:92–112

- Ahn G, Han S, Kim J, Kim S, Kim HB, Lee YS (2014) Empirical analysis of transport policy for regional development. The Korea Transport Institute, Sejong, (in Korean)
- Alkire S, Roche JM, Santos ME, Seth S (2011) Multidimensional poverty index 2011: brief methodological note. University of Oxford
- Andersen PL, Hansen MN (2012) Class and cultural capital—the case of class inequality in educational performance. *Eur Sociol Rev* 28(5):607–621
- Arthur D, Vassilvitskii S (2007) *k-means++*: the advantage of careful seeding. In: SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, 1027–1035. <https://dl.acm.org/doi/10.5555/1283383.1283494>
- Atkinson R (2006) Paddling the bunker: strategies of middle-class disaffiliation and colonization in the city. *Urban Stud* 43(4):819–832
- Atkinson R (2008) Commentary: gentrification, segregation and the vocabulary of affluent residential choice. *Urban Stud* 45:2626–2636
- Ballas D (2013) What makes a 'happy city'? *Cities* 32:39–50
- Bonaccorso G (2018) Mastering Machine Learning Algorithms: expert techniques to implement popular machine learning algorithms and fine-tune your models. Packt Publishing Ltd
- Byun S, Kim K (2010) Educational Inequality in South Korea: the widening socioeconomic gap in student achievement. In Emily H, Park H, Butler YG (eds) Globalization, changing demographics, and educational challenges in East Asia. Emerald Group Publishing Limited
- Byun S, Kim KK, Park H (2012) School choice and educational inequality in South Korea. *J School Choice* 6(2):158–183
- Chan TW, Goldthorpe JH (2007) Social stratification and cultural consumption: the visual arts in England. *Poetics* 35(2–3):168–190
- Chapple K (2007) Overcoming mismatch: beyond dispersal, mobility, and development strategies. *J Am Plan Assoc* 72(3):322–336
- Chen Y, Bouguerguene A, Li HX, Liu H, Shen Y, Al-Hussein M (2018) Spatial gaps in urban public transport supply and demand from the perspective of sustainability. *J Clean Prod* 195:1237–1248
- Choi EY (2004) The socio-economic segregation and the differentiation of public-sector schools. *J Urban Stud* 9:66–86. [in Korean]
- Choi EY (2006) The formation of rigid cycle of the rich in Gangnam—according to the change of Condo prices (1989–2004). *J Kor Urban Geogr Soc* 9(1):33–45. [in Korean]
- Choi MH, Sheong KS, Cho BM, Hwang IK, Chang HK, Kim MH, Hwang SS, Lim JS, Yoon TH (2011) Deprivation and mortality at the town level in Busan, Korea: an ecological study. *J Prev Med Public Health* 44(6):242–248
- Coleman JS (1990) Equality and achievement in education. Westview Press
- Cox S, Cox T (1996) Safety, systems and people. Butterworth Heinemann
- de la Espriella C (2009) Applications of poverty maps in urban planning: examples from Liberia, in Costa Rica. *Appl Spat Anal Policy* 3:163–182
- Duranton G, Puga D (2000) Diversity and specialisation in cities: why, where and when does it matter? *Urban Stud* 17(3):533–555
- George H (1973) Progress and Poverty; an inquiry into the cause of industrial depressions and of increase of want with increase of wealth: the remedy. AMS Press
- Grubestic TG, Wei R, Murray AT (2014) Spatial clustering overview and comparison: accuracy, sensitivity, and computational expense. *Ann Assoc Am Geogr* 104(6):1134–1156
- Grusky DB (2014) Social stratification: class, race, and gender in sociological perspective. Routledge
- Hacker KP, Seto KC, Costa F, Corburn J, Reis MG, Ko AI, Diuk-Wasser MA (2013) Urban slum structure: integrating socioeconomic and land cover data to model slum evolution in Salvador, Brazil. *Int J Health Geogr* 12:45
- Harvey D (1989) The urban experience. Johns Hopkins University
- Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference and prediction. Springer
- Henning C, Liao TF (2013) How to find an appropriate clustering for mixed-type variables with application to socio-economic stratification. *J R Stat Soc* 62:309–369
- Henning C (2015) What are the true clusters? *Pattern Recognit Lett* 64:53–62
- Hollingshead AB (1975) Four-factor index of social status. Yale University
- Hong E, Ahn BC (2011) Income-related health inequalities across regions in Korea. *Int J Equity Health* 10:41
- Horton FE, Reynolds DR (1971) Effects of urban spatial structure on individual behavior. *Econ Geogr* 47(1):36–48
- Im J, Hong SH (2017) Impact of a new subway line on housing values in Daegu, Korea: distance from existing lines. *Urban Stud* 55(15):3318–3335
- Jang M, Kang CD (2015) Housing prices in Seoul, Korea: a retail type and housing submarket approach. *Habitat Int* 49:516–528
- Johnstone B, Jones A, Kruger C (2019) Applied Unsupervised Learning with Python: discover hidden patterns and relationships in unstructured data with python. Packt Publishing Ltd, Birmingham
- Jung HS, Kim SW, Ahn SH (2014) Multidimensional Inequality in South Korea: an empirical analysis. *Asian Soc Work Policy Rev* 8(2):170–191

- Kang CD (2008) A social historical study on the Korean paranoid educational fever and status desire. *Kor Educ Rev* 14(2):5–32. [in Korean]
- Kang EK (2016) A study on wondosim urban regeneration of symbolic place and meaning of placeness through theater facilities. *Cineforum* 24:299–321. [in Korean]
- Kang HC (2014) Policy direction for decreasing the concentration of patients to extra-large hospitals. *Health Welf Policy Forum* 210:65–76. [in Korean]
- Kang HK (1991) Capital accumulation and the spatial division of classes: with special reference to the new middle class in Korea and Taiwan. *Korea J Popul Dev* 20(1):101–119
- Kaufman LB, Rousseeuw P (1990) Finding groups in data. Wiley
- KEIS (2019) Regional employment trends brief, 2019 Spring. Korea Employment Information Service [in Korean]
- Kernan JB, Bruce GD (1972) The socioeconomic structure of an urban area. *J Market Res* 9(1):15–18
- Kim CS (2014) The relationship between social security network and security life satisfaction in community residents: scale development and application of social security network. *J Korea Contents Assoc* 14(6):108–118. [in Korean]
- Kim DH, Park JA (2017) Spatial equity of neighborhood store accessibility: focused on the census on establishments (2006–2014) in Seoul, Korea. *J Korea Plan Assoc* 52(6):43–56. [in Korean]
- Kim E, Jeong YH (2003) Decomposition of regional income inequality in Korea. *Rev Reg Stud* 33(3):313–327
- Kim E, Kim E, Hong SW, Ha SJ (2003) Impacts of national development and decentralization policies on regional income disparity in Korea. *Ann Reg Sci* 37:79–91
- Kim JH (2007) A study on spatial equity of cultural facilities' distribution in Seoul. *J Geogr Educ* 51:43–59. [in Korean]
- Kingstone PW (2000) The classless society. Stanford University Press
- Lamont M, Fourier M (1992) Cultivating differences. Symbolic boundaries and the making inequalities. University of Chicago Press
- Lee BS, Chung EC, Kim YH (2002) The impacts of complex -specific characteristics on apartments' prices in Seoul. *Kukje Kyungje Yongu* 8(2):21–45. [in Korean]
- Lee CJ (2005) Korean education fever and private tutoring. *KEDI J Educ Policy* 2(1):99–107
- Lee JK (2009) Characteristics of the affordable housing based on disposable income of households in Seoul. *J Korea Plan Assoc* 44(7):97–108. [in Korean]
- Lee N, Rodri'guez-Pose A (2013) Innovation and spatial inequality in Europe and USA. *J Econ Geogr* 13:1–22
- Lelo K, Monni S, Tomassi F (2019) Socio-spatial inequalities and urban transformation. The case of Rime districts. *Socio-Econ Plan Sci* 68:1–11
- Levy F, Michel RC (1991) The economic future of American families: Income and wealth trends. Urban Institute Press, Washington
- Lin D, Allan A, Cui J (2015) The impacts of urban spatial structure and socioeconomic factors on patterns of commuting: a review. *Int J Urban Sci* 19(2):238–255
- Lucas K (2012) Transport and social exclusion: Where are we now? *Transp Policy* 20:105–113
- Marx K (1977) Capital: a critique of political economy. Vintage Books
- McLachlan G, Norman P (2020) Analyzing socio-economic change using a time comparable geodemographic classification: England and Wales, 1991–2011. *Appl Spat Anal Policy* 14:89–111
- McDowell L, Ward K, Perrons D, Ray K, Fagan C (2006) Place, class and local circuits of reproduction: exploring the social geography of middle-class children in London. *Urban Stud* 43(12):2163–2182
- Meltzer R, Schuetz J (2011) Bodegas or Bagel shops? Neighborhood differences in retail and household services. *Econ Dev Q* 26(1):73–94
- Monroe BL, Pan J, Roberts ME, Sen M, Sinclair B (2015) No formal theory, causal inference, and big data are not contradictory trends in political science. *Political Sci Politics* 48(1):71–74
- Murphy KP (2012) Machine learning: a probabilistic perspective. The MIT Press
- Mustard S, Ostendorf W (2005) Social exclusion, segregation and neighborhood effects. In: Kazepov Y (ed) Cities of Europe. Changing contexts, local arrangements and the challenge to urban cohesion. Blackwell
- Nijman J, Wei YD (2020) Urban inequalities in the 21st century economy. *Appl Geogr* 117:102188
- Noh E (2006) Statistical test of the regional income inequality in Korea. *Korean Econ Rev* 22(2):341–365
- Oh J (2017) South Korea's regional disparities and the 2018 Winter Olympics for regional development: Big Push revisited. *Int Area Stud Rev* 20(2):144–159
- Okulicz-Kozaryn A (2013) City Life: rankings (livability) versus perceptions (satisfaction). *Soc Indic Res* 110:433–451
- Panori A, Psycharis Y (2017) Exploring the links between education and income inequality at the Municipal Level in Greece. *Appl Spat Anal Policy* 12:101–126
- Park BG, Jang J (2020) The Gangnam-ization of Korean urban ideology. In: Doucette J, Park B-G (eds) Developmentalist cities? Interrogating urban developmentalism in East Asia. Haymarket Books
- Park T, Lee M, Han W (2015) KRIHS Policy Brief, 503. Korea Research Institute for Human Settlements [in Korean]
- Pekkanen J, Tuomilehto J, Uutela A, Vartiainen E, Nissinen A (1995) Social class, health behavior, and mortality among men and women in Eastern Finland. *BMJ Clin Res* 311(7005):589–593
- Piketty T (2014) Capital in the Twenty-First Century. The Harvard University Press
- Piketty T (2020) Capital and ideology. The Harvard University Press
- Pizzol B, Giannotti M, Tomasiello DB (2021) Qualifying accessibility to education to investigate spatial equity. *J Transp Geogr* 96:103199
- Rawls J (1971) A theory of Justice. Harvard University Press
- Reich R (1991) The work of nations. Vintage
- Roemer JE (1998) Theories of distributive justice. Harvard University Press
- Sen A (1992) Inequality reexamined. Harvard University Press
- Senlier N, Yildiz R, Aktas ED (2008) A perception survey for the evaluation of urban quality of life in Kocaeli and a comparison of the life satisfaction with the Europe cities. *Soc Indic Res* 94(2):213–226
- Seth MJ (2002) Education Fever; Society, Politics, and the Pursuit of Schooling in South Korea. University of Hawai'i Press
- Shevly E, Bell W (1955) Social area analysis: theory, illustrative application and computational procedures. Stanford University Press
- Shi Q, Dorling D (2020) Growing socio-spatial inequality in neo-liberal times? Comparing Beijing and London. *Appl Geogr* 115:102139
- Siqueira-Gay J, Giannotti M, Sester M (2019) Learning about spatial inequalities: capturing the heterogeneity in the urban environment. *J Clean Prod* 237:1–11
- Smith DM (1984) Alternative perspectives on 'urban inequality'. *Geoforum* 15(1):75–82
- Social Exclusion Unit (2003) Making the connections: final report on transport and social exclusion. Social Exclusion Unit
- Sohn J, Oh SK (2019) Explaining spatial distribution of the middle class: a multiple indicator approach with multiple explanatory dimensions. *Appl Spat Anal Policy* 12:871–905
- Soja EW (2010) Seeking spatial justice. University of Minnesota Press
- Son JJ (2004) Understanding Gangnam as a educational space. *Kor J Sociol Educ* 14(3):107–132
- Spector AN (1982) Towards integrating a model of urban socioeconomic structure and urban imagery. *Environ Plan A* 14:765–787
- Su S, Zhou H, Xu M, Ru H, Wang W, Weng M (2019) Auditing street walkability and associated social inequalities for planning implications. *J Transp Geogr* 74:62–76
- Talen E (2001) School, community, and spatial equity: an empirical investigation of access to elementary schools in West Virginia. *Ann Assoc Am Geogr* 91(3):465–486
- Talley W (1996) Linkages between transportation infrastructure investment and economic production. *Logist Transp Rev* 32(1):145–154
- Thobcke E, Charumilind (2002) Economic Inequality and Its Socioeconomic Impact. *World Development* 30(9):1477–1495
- van Kempen R (2005) Segregation and housing conditions of immigrants in western European cities. In: Kazepov Y (ed) Cities of Europe. Changing contexts, local arrangements and the challenge to urban cohesion. Blackwell
- Veblen T (1994) The theory of the leisure class. Penguin Books
- Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer
- von Luxburg U, Williamson RC, Guyon I (2012) Clustering: Science or art? In: Guyon I, Dror G, Lemaire V, Taylor G, Silver D (eds) JMLR: Workshop and Conference Proceedings, Vol. 27. pp. 65–79. <http://proceedings.mlr.press/v27/luxburg12a/luxburg12a.pdf>
- Wright EO (2005) Approaches to class analysis. Cambridge University Press
- Wu Y, Wei YD, Li H (2020) Analyzing spatial heterogeneity of housing prices using large datasets. *Appl Spat Anal Policy* 13:223–256
- Yang H, Kang E, Kim MG, Koh K (2020) Is there regional inequality in the medical accessibility of the severely injured? Application of driving time data in South Korea. *J Market Econ* 49(1):1–29. [in Korean]
- Yang M (2018) The rise of 'Gangnam style': manufacturing the urban middle classes in Seoul, 1976–1996. *Urban Stud* 55(15):3404–3420
- Yea S (2000) Maps of resistance and geographies of dissent in the Cholla Region of South Korea. *Korean Stud* 24:69–93
- Yi C, Kang YE, Kim MW, Kim HW (2012) Efficient utilization of existing parking spaces in Seoul. The Seoul Institute [in Korean]
- Zamboni I, Serra P, Sauri D, Carlucci M, Salvati L (2017) Beyond the 'Mediterranean city': socioeconomic disparities and urban sprawl in three Southern European cities. *Geogr Ann Ser B* 99(3):319–337
- Zhang X, Kanbur R (2005) Spatial inequality in education and health care in China. *China Econ Rev* 16(2):189–204

Zhang SS (1998) A study on the household's socio-economic characteristics affecting the choice of apartment. J Archit Inst Korea 14(11):31–38. [in Korean]

Competing interests

The author declares no competing interests.

Ethical approval

Not applicable.

Informed consent

Not applicable.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-022-01035-5>.

Correspondence and requests for materials should be addressed to Seungwoo Han.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022