Revealing Urban Activity Patterns Around Metro Stations through Social Media Network Data

Sanaz Mohammadbagherzadeh (Mohammadbagherzade19@itu.edu.tr)
Istanbul Technical University
Fatih Terzi
Istanbul Technical University

Research Article

Keywords: Metro station, urban activity, social media network data, spatial analysis, Istanbul

Posted Date: April 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1549279/v1

Abstract

Stations on the metro system are important transportation hubs that handle a large number of the daily trip. Through the metro stations that facilitate pedestrian mobility and serve as a magnet for businesses, public transportation systems have the potential to alter the character of a community depending on the circumstances. Using location-based social network (LBSN) data, this study aims to quantify the area surrounding metro stations and metro lines, as well as to conduct a cluster analysis of Istanbul metro stations based on the pattern of LBSN data to gain insights into the characteristics of each station's surrounding area. To this end, the Foursquare venue data were gathered using the Foursquare API database. Then the data was cleaned and categorized based on the main activity and mapped in ArcGIS pro. K means cluster analysis was performed based on different venue activities around each station. Based on various venue activities, 77 Metro stations are classified into three types, namely undeveloped stations, developed stations, and touristic stations. The majority of stations were placed in undeveloped groups with a low activity level. The developed stations are mainly located in the commercial and touristic areas.

1. Introduction

The extension of urban metro system network has been a crucial strategy to resolve the rising issue of traffic congestion and make wise consumption of the limited resources in big cities, helping to achieve sustainable development in urban transportation. In this system, stations, as transport nodes, serve a critical role in handling a large number of intra-urban journeys, providing accessibility to their immediate surroundings [1]. As networks' hubs, stations are often seen as nodes of socio-economic activity due to their role in generating mobility [2] which has the potential to create an accumulation impact for a variety of everyday activities that are diverse, dense, and walkable [3]. Consequently, the catchment area of the metro station attracts a significant proportion of human activities and urban functions, transforming metro stations into lively and dynamic urban daily life areas [4], which have become the lifeblood of local communities in many forms, eventually satisfying the social, environmental, and economic demands of people [5]. Therefore, due to the accessibility and potential of the catchment area of stations as a concentrated location for urban activity, they are appropriate locations for developing structured nodes in urban networks [6]. Also, the authors in [7] discussed that revitalizing the service areas of metro stations for urban regeneration plays a substantial role. They mentioned that as the metro is the foundation of transportation infrastructure, increasing residential and commercial land will increase metro liveliness. Meanwhile, another study shows that land-use patterns in a region straightly affect transportation usage and, therefore, the long-term sustainable usage of mass transport systems [8].

In this context, adopting proper quantification approaches to observe and analyze the operation and development situation of metro stations' catchment areas will aid urban planners in understanding and dealing with the pattern of land development and spatial form of surrounding metro stations [9]. One of the main analytical methods is the classification of transit nodes. Such a categorization, in particular, takes the diversity of individual stations and analyzes the similarities within clusters of nodes, enabling planners and governors to formulate more focused strategies for development [10].

Various researchers have tried to categorize stations based on surrounding area characteristics [11–12–13, 14]. Existing research relies primarily on conventional data such as censuses, surveys, and land use. However, collecting and regularly updating this data for a large number of sites is expensive and time-consuming [15]. Due to the rise of big data and the mobile internet, novel data types have been increasingly explored and used in urban studies, giving researchers fresh insight into the time and spatial dimensions [9]. In [16], the authors discovered that urban analytics and big data increase our knowledge of urban systems by generating fresh and unique hypotheses quicker than before.

The purpose of this paper is to investigate the spatial patterns of data emerging from location-based social media networks (LBSN) in order to quantify the area around metro stations and metro lines, as well as to conduct a cluster analysis of Istanbul metro stations based on the pattern of LBSN data in order to gain insights into the station area's characteristics. The paper concentrates on the following research questions:

- Is there any cluster pattern of metro stations based on location-based social network (LBSN) data?
- To what extent are there differences in terms of urban activities around metro stations?

Urban Activity and Metro Stations

Most existing research on metro station areas stems from the transit-oriented development (TOD) field. Transport Oriented Development (TOD) is considered a strategy to integrate transportation systems and land use by concentrating the urban development within the catchment area of transit nodes [17]. TOD's main objective is to create a compact urban form with mixed land use (residential, work, open space, entertainment, and leisure facilities) and to develop a pedestrian- and cycling-friendly environment within walking distance (e.g., 400–800) of public transport nodes [18]. These characteristics are often described as the '3Ds': high-density development, land use diversity, and effective urban design [19]. Consequently, this encouraging development has resulted in a dense built environment immediately around transit stations, attracting various commercial businesses, pedestrian activities, and economic advantages. In addition to providing efficient transportation services, the catchment area of transit nodes such as metro stations is a critical element of social spaces, enabling human activities and facilities to live, entertain, and socialize within the station area. Therefore, it is necessary to analyze station areas' social functions and urban activities thoroughly.

In [5], the authors studied the distribution of human activities within the metro station's catchment area using social media data from Facebook and concluded that metro TOD zones significantly support Hong Kong people's social, recreational, and economic demands. Another study [4] classified metro stations using points of interest (POIs) in Hong Kong. Results from the analysis of urban functions indicate four themes strongly linked to commercial and residential activities and tourist and industrial activities. There is a noticeable spatial concentration of metro stations that fall into the same thematic category, implying that they provide the same urban functions. Using smart card data, [20] examined the Wuhan metro network and explored the similarities and contrasts between land use distribution and the functions surrounding the station area. Researchers were able to determine each station's catchment area's function by analyzing human activities distribution and characteristics around each metro station.

Location Based Social Network

An increasing number of smartphone users has resulted in the development of a range of applications, the most prominent of which are social networking services (SNS) applications like Facebook and Twitter. Location-based social networks (LBSN) are social networking services that leverage global positioning system (GPS) location data, such as Facebook and Foursquare [21]. LBSNs produce information about places in social networks and a social structure formed by places through which members may exchange thoughts, evaluate suggestions, and develop knowledge about shared interests, behaviors, and activities based on people's location-tagged information [22]. Therefore, it offers a fresh look into how individuals like to spend their time in cities, both in terms of location and duration [23], and these footprints provide evidence of people's choices and patterns of use in specific urban spaces resulting in the emergence of a new generation of human knowledge about cities [24].

Location-based social networks (LBSN) were divided into three main categories: Photo/video sharing (e.g., Instagram, Flickr, YouTube), Microblogs (e.g., Twitter, Facebook, Weibo), and Point-of-Interest (POI) (e.g., Foursquare, Open Street Map, Yelp) [25]. Two new methods for analyzing urban areas are points of interest and social media check-in data [26]. By evaluating the ratio and frequency of POI visits across locations, the POI application's data was most commonly utilized to track activity patterns of the site and get information on how people feel about certain areas [25]. The POI data indicates the primary land use of various land parcels and the primary function of various structures or complexes. Despite its more conventional equivalents (for instance, an official land use map), revised and released once or twice a year, LBSN continually monitors and updates the POI data [27]. Substantial information about land use is documented in social media, both directly and indirectly. Through the characterization of activities, indirect information can be obtained. Also, the POI check-ins provide precise land use information [28].

A variety of urban studies have used POI from social media to augment their findings. In [26], the authors evaluated data obtained from location-based social networks to study the area's activity patterns of Istanbul's historic districts. Another study investigated the uses of LBSN data by examining the relationship between Foursquare venues data of Seoul produced via user engagement and the urban's features [29]. The paper [30] performed an analysis using Foursquare venues data to examine the qualitative and quantitative spatial allocation of coffee shops in Seattle. In [15], the authors used non-traditional data such as social media check-ins to supplement traditional data like censuses and interviews to improve the quality of research and monitoring of transit-served areas.

Authors in [31] utilized the Foursquare social network's data to illustrate Murcia's commercial activity. The resulting visualization shows activity patterns in Murcia as well as cities characteristics based on individuals’ preferences and interests. In other research, they utilized location-based social media network data to recognize effective public places [32].

2. The Study Area: Istanbul

Istanbul is located on both sides of the Bosporus strait, straddling Europe and Asia. It has a population of 16 million people, accounting for 18.71 percent of Turkey's total population [33], and is both Europe's and Turkey's most populated metropolis. Istanbul, a fast expanding city, is confronted with substantial urban issues such as major urban transport problems leading to persistent traffic congestion, car crashes, air pollution and noise pollution, and energy and time waste [34]. In [35], the authors stated that traffic congestion, particularly the Istanbul Strait crossings, is one of the city's most persistent difficulties, according to social studies.

Annual motorization rates in Turkey are increasing at a rate of 4.5 percent per year, in accordance with economic and demographic expansion [36]. As a solution to this problem, Istanbul Metropolitan Municipality (IMM), responsible for Istanbul's urban mobility planning, has devoted a significant portion of its funds to expanding public transportation capacity over the previous decade. One of these projects is to expand the rail system network aiming to make it the city's transportation system's backbone.

Between 2004 and 2018, IMM extended the city rail network from 45.1 kilometers to 158 kilometers. IMM intends to significantly expand the rail system of Istanbul to 495.35 kilometers by 2019 and 710.95 kilometers by 2023 [36].

The extension of Istanbul's metro system is transforming the city's transportation infrastructure and significantly influences the city's urban development [37].

In [38], the authors studied the spatial impact of the Istanbul Metro network by considering a 2 km buffer around it. They state that the existing rail system serves 46% of the population. While this amount will increase to 71% and 87% in 2019 and 2023, respectively [38]. Paper [39] mentioned that the Istanbul Metro provided indirect economic advantages, particularly in corporate investment and industry shifts from industrial and manufacturing-related industries to high-profile service-based enterprises within its catchment regions.

The Istanbul Metro is a rapid transit railway network that serves Istanbul's European and Asian (Anatolian) sides. The system consists of eight lines and 107 stations. Also, the Istanbul Metro is connected to the underground systems of the Marmaray. The Marmaray Mass Rapid Transit railway system connects the railway tracks on both sides of Istanbul through a tunnel under the Istanbul Strait. It has 43 stations and starts from Halkalı on the European side and ends in the Gebze district of Kocaeli. The study area (Fig. 1) covers the metro and Marmaray systems, located on the European side of Istanbul. It consists of seven lines and 77 stations. The three main metro lines of Istanbul, shown in Fig. 2, were built in different periods. The starting and destination points of these lines; socio-economic development of the neighborhoods where the metro stops are connected along the line; and land use patterns also differ. Undoubtedly, this differentiation causes the rates of activity types that fall under the influence of metro areas to differ.

The Hacıosman-Yenikapı metro line (Yenikapı, Vezneciler, Şişhane, Taksim, Osmanbey, Haliç, Şişli, Gayrettepe, Levent); Hacıosman, the starting station, is connected to the Central Business District of Istanbul, and Yenikapı station is connected to the intercity passenger port. Most of the stops on the line are connected with the central business district (Taksim, Osmanbey, Haliç, Şişli, Gayrettepe, Levent) and white-collar workplaces, including mixed-use, and establish a relationship with the Historic Peninsula. In addition, the world-famous Taksim Square is also located on this line.

The Yenikapı-Bağcılar-Şirinevler metro line (Aksaray, Bakırkoy, Ataköy, Yenibosna); is one of the oldest metro lines, connecting with the starting station, Sirkeci Historical Peninsula. The destination stop is linked to Yenibosna and the former airport (Atatürk Airport). Some of the stops on this line are connected with old neighborhoods (some of them transformed from squatter areas into apartments, such as Terazidere, Bağcılar, Güngören).

The Marmaray line (Sirkeci) is parallel to the Marmara coast, connecting the Anatolian and European Sides. The line, which starts with the Sirkeci stop on the European side, ends with Küçükçekmece, which is one of the highly populated residential areas.

3. Data And Methodology

The venue information data is extracted from the location-based social network (LBSN) Foursquare. Foursquare is one of the most prominent LBSNs allows members to "check-in" to a venue to express impressions and publicize their visit. It is also a social suggestion tool that enables users to search for and identify the best areas based on the activities of others. According to the sort of activity that may be carried out at a venue, Foursquare categorizes urban activities —venues— into ten basic groups. Foursquare's ten primary categories venues are Art & Entertainment, College & Education, Events, Food, Nightlife Spots, Outdoors & Recreation, Professional & Other Places, Residences, Shop & Services, and Travel & Transport.

The methodology is divided into three distinct steps (Fig. 2)

- Data Collection: Extracting the data using Foursquare's Application Programming Interface (API).

- Kernel Density Analysis: Investigating density distribution venue data.
- K-Mean Cluster Analysis: Evaluating and categorizing station areas based on the pattern of venues around them.

3.1. Data Collection

We used Python to write the code based on Foursquare's API. Several points were selected in the study area. The 500-meter buffer zones around seven metro lines and the location attributes of these points with a 500-meter search radius were considered as Foursquare's venue search API. For the response, 7162 venue records were collected and converted to a CSV text file (Fig. 2). The venue points dataset includes venue location (latitude and longitude), venue name, and venue category.

Overall characteristics and venue distribution are shown in Table 1. The categories with the most significant numbers of venues were Food (~ 51%), Shop and Service (~ 18%), and Outdoors & Recreation (~ 12%).

Table 1
Distribution of venues according to the primary category

Category Name	Number of Venues	%
Food	3648	50.93
Shop & Service	1261	17.61
Outdoors & Recreation	865	12.08
Arts & Entertainment	497	6.93
Travel & Transport	403	5.63
Nightlife Spot	360	5.03
Professional & Other Places	106	1.48
College & University	15	0.21
Event	4	0.06
Residence	3	0.04
Total	7162	100.00

Since the number of venues for college and education, events, and residences was relatively low in the study area, they were eliminated from the database. As shown in Fig. 2, data cleaning is performed after collecting data using Python. It entails translating the Python program's text data to GIS data. Cleaned data that contains venue points with information about the location (latitude and longitude) is mapped in ArcGIS Pro as point shapefile (Fig. 3).

3.2. Kernel Density Analysis

Kernel density utilizes to show the point's density distribution smoothly based on the chosen distance [40]. For kernel density analysis, similar categories according to their function were merged into three main groups (Table.2).

Table 2
Distribution of venues according to merged category

Category Name	Number of Venues	%
Food & Beverage	4008	0.56
Shop & Service	1770	0.25
Outdoors & Entertainment	1362	0.19
Total	7140	

In this step, the collected venues are demonstrated as point features, and by using ArcGIS Pro, kernel density estimation is used to analyze their density's spatial distribution (Fig. 2). ArcGIS Pro's kernel density tool was applied to Food & Beverage, Shop & Service, Outdoors & Entertainment, and Total Venue to investigate spatial patterns of different categories. Raster maps were constructed, with continuous surfaces representing the intensity or severity of point features. The severity of point features is represented by color codes (Fig. 6).

3.3. K-Mean Cluster Analysis

Clustering categorizes a data set based on the degree of similarity within the same group and the degree of dissimilarity across the group [41]. In this study, a K-means algorithm was applied to classify each of the 77 stations into groups based on the venue characteristics located around each (Fig. 2). In cluster analysis, the K-means clustering is the most well-known and oldest approach. It has been extensively researched, with numerous expansions, and used in a broad range of fundamental fields [42]. For station cluster analysis, the main Foursquare categories were utilized to investigate the activities around stations more precisely. Using the spatial join tool in ArcGIS Pro, the venues within 500 meters of each metro station were collected (Fig. 4), and the number of venues for each category was counted (Table.3).

Table 3
Part of distribution of venue around stations

Station Name	Travel & Transport	Shop & Service	Outdoors & Recreation	Food	Nightlife Spot	Professional & Other Places	Arts & Entertainment	Total
4.levent	1	4	9	34	6	0	3	57
Aksaray	6	5	5	49	4	4	1	74
Ataköy	4	8	9	24	3	0	0	48

By applying the elbow method, the optimal number of clusters (k) is defined as three. Cluster analysis results are plotted in a boxplot, and each cluster's statistics are represented (Fig. 5). Also, the average number of venues for each category was listed in Table 4. Meanwhile, for better visualization, a color was assigned to each cluster mapped in ArcGIS Pro (Fig. 7).

Table 4
Statistics of each cluster

Cluster	Cluster Size	%	Travel & Transport	Shop & Service	Outdoors & Recreation	Food	Nightlife Spot	Professional & Other Places	Arts & Entertainment	Total
0	59	76.62%	4.25	10.83	7.29	32.53	3.56	0.92	4.03	63.42
1	14	18.18%	18.14	43.43	29.57	133.93	15.57	4.07	16.5	261.29
2	4	5.9%	27	26.75	13.75	74.75	13.25	4.25	17.5	177.25

4. Results And Evaluation

The study's findings are provided within the context of the research questions posed in the first part.

The distribution of venues around the Istanbul metro station

The outcomes of kernel density analysis (Fig. 6) show that total venues form a significant density along the Hacıosman-Yenikapı metro line (Yenikapi, Vezneciler, Haliç, Şişhane, Taksim, Osmanbey, Şişli, Gayrettepe, Levent) and Marmaray line (Sirkeci, Zeytinburnu, Bakirköy) and Yenikapı-Bağcılar metro line (Aksaray, Sagmalcilar, Bagcilar Meydan, Bakirkoy, Atakoy, Yenibosna).

Food & beverage venues form significant density along the Hacıosman-Yenikapı metro line (Yenikapı, Vezneciler, Şişhane, Taksim, Osmanbey, Haliç, Şişli, Gayrettepe, Levent), the Marmaray line (Sirkeci), and the Yenikapı-Bağcılar-Şirinevler metro line (Aksaray, Bakırkoy, Ataköy, Yenibosna).

Shop & service venues form significant densities along the Hacosman-Yenikapı metro line (Yenikapi, Vezneciler, Şişhane, Taksim, Osmanbey, Haliç, Şişli, Gayrettepe, Levent) and the Marmaray line (Sirkeci), the Yenikapı-Bağcılar-Şirinevler metro line (Aksaray) and the Kirazlı-Bağcılar metro line (İSTOÇ).

Outdoors & entertainment venues form significant densities along the Hacıosman-Yenikapı metro line (Yenikapi, Vezneciler, Haliç, Şişhane, Taksim, Osmanbey, Şişli, Gayrettepe, Levent, 4. Levent), the Marmaray line (Sirkeci, Bakirkoy, Atakoy) and the Yenikapı-Bağcılar-Şirinevler metro line (Sagmalcilar, Otogar, Esenler, Bagcilar Meydan, Zeytinburnu).

The analysis significantly reveals that the density distribution of all three venue categories along the Hacıosman-Yenikapı metro line from Yenikapi station to Levent station was high. While the density distribution of three venue categories along the Marmaray line and the Yenikapı-Bağcılar-Şirinevler metro line was different. The newly built lines, the Mahmutbey-Mecidiyeköy line and the Bahariye-Olimpiyat line have significantly lower density distribution, while the Kirazlı-Bağcılar-Metrokent line, which has been active since 2013, has a similar situation to newly built lines.

Spatial Clustering of Urban Activity Around Metro Stations

Three groupings of stations emerged from the cluster analysis. "Cluster 0" contains 59 stations which are almost 75% of all stations. It has a lower venue in all categories than "cluster 1" and "cluster 2" (Fig. 7). In "Cluster 0", outdoors & recreation and food &beverages have higher venues than any other category. Figure 7 reveals that all stations of all lines studied except the Hacıosman-Yenikapı metro line were grouped in the "Cluster 0".

"Cluster 1" has 14 stations which are almost 18% of all stations. The distribution of the venues around stations of "cluster 1" is high. In comparison to other categories Shop & Service, Outdoors & Recreation, and Food have a higher number of venues. It can be

observed that the majority of these stations are located on the Hacıosman-Yenikapı metro line (Osmanbey, Şişli, Levent, 4. Levent). There are 2 stations from Levent-Boğaziçi Üniversitesi metro line (Nispetiye, Boğaziçi) and 5 stations in Yenikapı-Bağcılar-Şirinevler metro line (Aksaray, Sagmalcilar, Bagcilar Meydan, Bakirkoy, Atakoy) and two stations in Marmaray line (Zeytinburnu, Bakirkoy) and a station in newly opened Mahmutbey-Mecidiyeköy metro line (Veysel Karani Akşemsettin).

"Cluster 2" has four stations which are almost 6% of all stations. These categories of stations have the highest concentration of venues around them. It can be seen that the categories of Travel & Transport, Professional & Other Places, and Arts & Entertainment outnumber other clusters. Meanwhile, the Outdoors & Recreation category has the lowest number of venues among other clusters; Vezneciler, Şişhane, and Taksim stations of the Hacıosman-Yenikapı metro line and Sirkeci of the Marmaray line were categorized in this cluster.

Cluster analysis results reveal an overview of the area's characteristics around stations. The stations in "Cluster 0" have the lowest number of venues representing undeveloped station areas. While stations' area in "Cluster 1" has a high number of venues, and it can be named as developed stations' area. However, four stations grouped in "cluster 2" have very different characteristics in terms of categories and number of venues. Since these four stations, Vezneciler, Şişhane, and Taksim and Sirkeci, are located in Istanbul's historical and touristic area, their difference from the developed stations' area can be explained. So this cluster can be named as touristic stations' area.

5. Conclusion And Discussion

Investment in public transportation systems has the potential to alter the character of a community temporarily or permanently through the stations that facilitate pedestrian mobility and serve as a magnet for businesses. As a result, it becomes critical to present a Metro Railway typology, which classifies each existing and planned station in metro networks according to its unique characteristics. One of these characterization methods could be to investigate work and activity spaces in order to gain a better understanding of the urban life that develops around metro stations.

The majority of station-related research relies on traditional data such as censuses, surveys, and land use, despite the fact that collecting and updating this data for many sites would be costly and time-consuming. With the rise of big data and mobile internet, new data types are being explored and used more frequently in urban studies.

The purpose of this paper is to assess the spatial patterns of data extracted from location-based social media networks (LBSN) in order to quantify the area surrounding metro stations and lines and to conduct a cluster analysis of Istanbul metro stations based on the LBSN data pattern in order to gain insight into the station area's characteristics.

Kernel density analysis was applied by ArcGIS pro to investigate the density distribution of venues generated by Foursquare along the Istanbul metro lines. The outcomes show that other metro lines have different venues for each category, except for the Hacıosman-

Yenikapı metro line, with the highest venues density in all categories. It can be assumed that since the Hacıosman-Yenikapı metro line passes through historical and commercial parts of Istanbul, the concentration of venues is high.

Then, the k-means algorithm was performed to classify each of the 77 stations into groups based on the venue characteristics located around each of them.

Cluster analysis results reveal an overview of the area's characteristics around stations. The stations' areas in "Cluster 0" have the lowest number of venues representing undeveloped stations areas. While the stations' areas in "Cluster 1" have a high number of venues, they can be named "developed stations' areas." However, stations grouped into "cluster 2" have very different characteristics in terms of categories and the number of venues. Since these stations are located in Istanbul's historical and touristic areas, their difference from developed stations' areas can be explained. So this cluster was named the "tourist stations' area."

This classification helps planners and policymakers to formulate more focused strategies for development around stations. The characteristics of the developed stations can be used as indicators of how urban spaces around undeveloped stations can undergo development.

Declarations

Ethics approval and Consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data generated or analyzed during this study are included in the paper.

Competing interests

The authors declare that they have no conflicts of interest.

Funding

Not applicable

Authors' contributions

The authors of the paper -Sanaz Mohammadbagherzadeh and Fatih Terzi- contributed to all parts of the paper.

GIS analyses, maps, and tables were carried out by Sanaz Mohammadbagherzadeh.

All authors reviewed the manuscript.

Acknowledgements

Not applicable

References

1. Gong Y, Lin Y, Duan Z (2017) Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. Computers Environ Urban Syst 64:169–183. https://doi.org/10.1016/j.compenvurbsys.2017.02.003
2. Bernick M, Cervero R (1997) Transit villages in the 21st century. McGraw-Hill 95:387
3. Loo BP, Cheng AH, Nichols SL (2017) Transit-oriented development on greenfield versus infill sites: Some lessons from Hong Kong. Landsc Urban Plann 167:37–48. https://doi.org/10.1016/j.landurbplan.2017.05.013

4. Yu Z, Zhu X, Liu X (2022) Characterizing metro stations via urban function: Thematic evidence from transit-oriented development (TOD) in Hong Kong. Journal of Transport Geography. 1;99:103299. https://doi.org/10.1016/j.jtrangeo.2022.103299
5. Lang W, Hui EC, Chen T, Li X (2020) Understanding livable dense urban form for social activities in transit-oriented development through human-scale measurements. Habitat Int 1:104:102238. https://doi.org/10.1016/j.habitatint.2020.102238
6. Bertolini L, Spit T (2005) Cities on rails: The redevelopment of railway stations and their surroundings. Routledge
7. Tu W, Zhu T, Zhong C, Zhang X, Xu Y, Li Q (2021) Exploring metro vibrancy and its relationship with built environment: a cross-city comparison using multi-source urban data. Geo-spatial Inform Sci 2:1–5. https://doi.org/10.1080/10095020.2021.1996212
8. AlKhereibi A, AlSuwaidi M, Al-Mohammed R, Pokharel S, Ayari M (2021) An integrated urban-transport smart growth model around metro stations: A case of Qatar. Transp Res Interdisciplinary Perspect. https://doi.org/10.1016/j.trip.2021.100392. .1;10:100392
9. Ye Z, Chen Y, Zhang L (2017) The analysis of space use around Shanghai metro stations using dynamic data from mobile applications. Transportation research procedia.1;25:3147-60. https://doi.org/10.1016/j.trpro.2017.05.353
10. Papa E, Pagliara F, Bertolini L (2008) Rail system development and urban transformations: Towards a spatial decision support system. In: Bruinsma, F., Pels, E., Rietveld, P., Priemus, H., van Wee, B. (eds) Railway Development. Physica-Verlag HD. https://doi.org/10.1007/978-3-7908-1972-4_16
11. Reusser DE, Loukopoulos P, Stauffacher M, Scholz RW (2008) Classifying railway stations for sustainable transitions–balancing node and place functions. Journal of transport geography. 1;16(3):191–202. https://doi.org/10.1016/j.jtrangeo.2007.05.004
12. Chen X, Lin L (2015) The node-place analysis on the “hubtropolis” urban form: The case of Shanghai Hongqiao air-rail hub. Habitat Int. https://doi.org/10.1016/j.habitatint.2015.06.013. 1;49:445 – 53
13. Monajem S, Nosratian FE (2015) The evaluation of the spatial integration of station areas via the node place model; an application to subway station areas in Tehran. Transp Res Part D: Transp Environ 1:40:14–27. https://doi.org/10.1016/j.trd.2015.07.009
14. Groenendijk L, Rezaei J, Correia G (2018) Incorporating the travellers’ experience value in assessing the quality of transit nodes: A Rotterdam case study. Case Studies on Transport Policy. 1;6(4):564 – 76. https://doi.org/10.1016/j.cstp.2018.07.007
15. Zhou J, Wang Q, Liu H (2018) Evaluating transit-served areas with non-traditional data. Journal of Transport and Land Use. 1;11(1):1323-49
16. Kandt J, Batty M (2021) Smart cities, big data and urban policy: Towards urban analytics for the long run. Cities. 1;109:102992. https://doi.org/10.1016/j.cities.2020.102992
17. Cervero R (2004) Transit-oriented development in the United States: Experiences, challenges, and prospects. Transportation Research Board
18. Calthorpe P (1993) The next American metropolis: Ecology, community, and the American dream. Princeton architectural press
19. Cervero R, Kockelman K (1997) Travel demand and the 3Ds: Density, diversity, and design. Transportation research part D: Transport and environment. 1;2(3):199–219. https://doi.org/10.1016/S1361-9209(97)00009-6
20. Zhou Y, Fang Z, Zhan Q, Huang Y, Fu X (2017) Inferring social functions available in the metro station area from passengers’ staying activities in smart card data. ISPRS Int J Geo-Information 6(12):394. https://doi.org/10.3390/ijgi6120394
21. Allamanis M, Scellato S, Mascolo C (2012) Evolution of a location-based online social network: analysis and models. InProceedings of the 2012 Internet Measurement Conference 14:145–158. https://doi.org/10.1145/2398776.2398793
22. Zheng Y (2011) Location-Based Social Networks: Users. In: Zheng Y, Zhou X (eds) Computing with Spatial Trajectories. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-1629-6_8
23. Shen Y, Karimi K (2016) Urban function connectivity: Characterisation of functional urban streets with social media check-in data. Cities 1 55:9–21. https://doi.org/10.1016/j.cities.2016.03.013
24. Wyly E (2014) The new quantitative revolution. Dialogues in Human Geography 4(1):26–38. https://doi.org/10.1177/2043820614525732
25. McKitrick MK, Schuurman N, Crooks VA (2022) Collecting, analyzing, and visualizing location-based social media data: review of methods in GIS-social media analysis. GeoJournal 19:1–23. https://doi.org/10.1007/s10708-022-10584-w
26. Üsküplü T, Terzi F, Kartal H (2020) Discovering activity patterns in the city by social media network data: a case study of Istanbul. Appl Spat Anal Policy 13(4):945–958. https://doi.org/10.1007/s12061-020-09336-5
27. Zhou J, Yang Y, Gu P, Yin L, Zhang F, Zhang F, Li D (2019) Can TODness improve (expected) performances of TODs? An exploration facilitated by non-traditional data. Transp Res Part D: Transp Environ 1:74:28–47. https://doi.org/10.1016/j.trd.2019.07.008

28. Yang Y, Heppenstall A, Turner A, Comber A (2019) Who, where, why and when? Using smart card and social media data to understand urban mobility. ISPRS Int J Geo-Information 8(6):271. https://doi.org/10.3390/ijgi8060271
29. Hong I (2015) Spatial analysis of location-based social networks in seoul, korea. J Geographic Inform Syst 20(03):259
30. Hong I, Jung JK (2017) What is so “hot” in heatmap? Qualitative code cluster analysis with foursquare venue. Cartographica: The International Journal for Geographic Information and Geovisualization 52(4):332–348. https://doi.org/10.3138/cart.52.4.2016-0005
31. Agryzkov T, Martí P, Tortosa L, Vicent JF (2017) Measuring urban activities using Foursquare data and network analysis: a case study of Murcia (Spain). Int J Geogr Inf Sci 2(1):100–121. https://doi.org/10.1080/13658816.2016.1188931
32. Martí P, Serrano-Estrada L, Nolasco-Cirugeda A (2017) Using locative social media and urban cartographies to identify and locate successful urban plazas. Cities. 1;64:66–78. https://doi.org/10.1016/j.cities.2017.02.007
33. TurkStat (2022) Turkish statistical institute. Online Data Base. https://data.tuik.gov.tr/Bulten/Index?p=The-Results-of-Address-Based-Population-Registration-System-2021-45500&dil=2. Accessed Feb 2022
34. IMM (2015) Strategic Plan for 2015–2019. Retrieved from http://www.ibb.gov.tr/en-US/Organization/Birimler/StratejikPlanlamaMd/Documents/stratejik_plan_2015–2019.pdf
35. Alpkokin P, Black JA, Iyinam S, Kesten AS (2016) Historical analysis of economic, social and environmental impacts of the Europe-Asia crossings in Istanbul. Int J Sustainable Transp 10(2):65–75. https://doi.org/10.1080/15568318.2013.853852
36. IMM (2016) Transport Annual Report. Retrieved from https://tuhim.ibb.gov.tr/media/ 2131/imm_transport_report.pdf
37. Canitez F, Alpkokin P, Kiremitci ST (2020) Sustainable urban mobility in Istanbul: Challenges and prospects. Case studies on transport policy.1;8(4):1148-57. https://doi.org/10.1016/j.cstp.2020.07.005
38. ÖNDEN İ, ÇAKMAK E, İSTANBUL’UN ARAZİ KULLANIM DEĞİŞİMİ VE METRO AĞI KAPSAMA ALANLARINA YÖNELİK MEKANSAL ANALİZLER (2020) Beykoz Akademi Dergisi 8(2):328–340
39. Beyazit E (2015) Are wider economic impacts of transport infrastructures always beneficial? Impacts of the Istanbul Metro on the generation of spatio-economic inequalities. J Transp Geogr 1:45:12–23. https://doi.org/10.1016/j.jtrangeo.2015.03.009
40. Chainey S, Reid S, Stuart N (2002) When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime. Socio-economic Appl geographic Inform Sci 5:9:22–36
41. Kaufman L, Rousseeuw PJ (2009) Finding groups in data: an introduction to cluster analysis. John Wiley & Sons
42. Yang MS, Sinaga KP (2019) A feature-reduction multi-view k-means clustering algorithm. IEEE Access 7:114472–114486

Figures

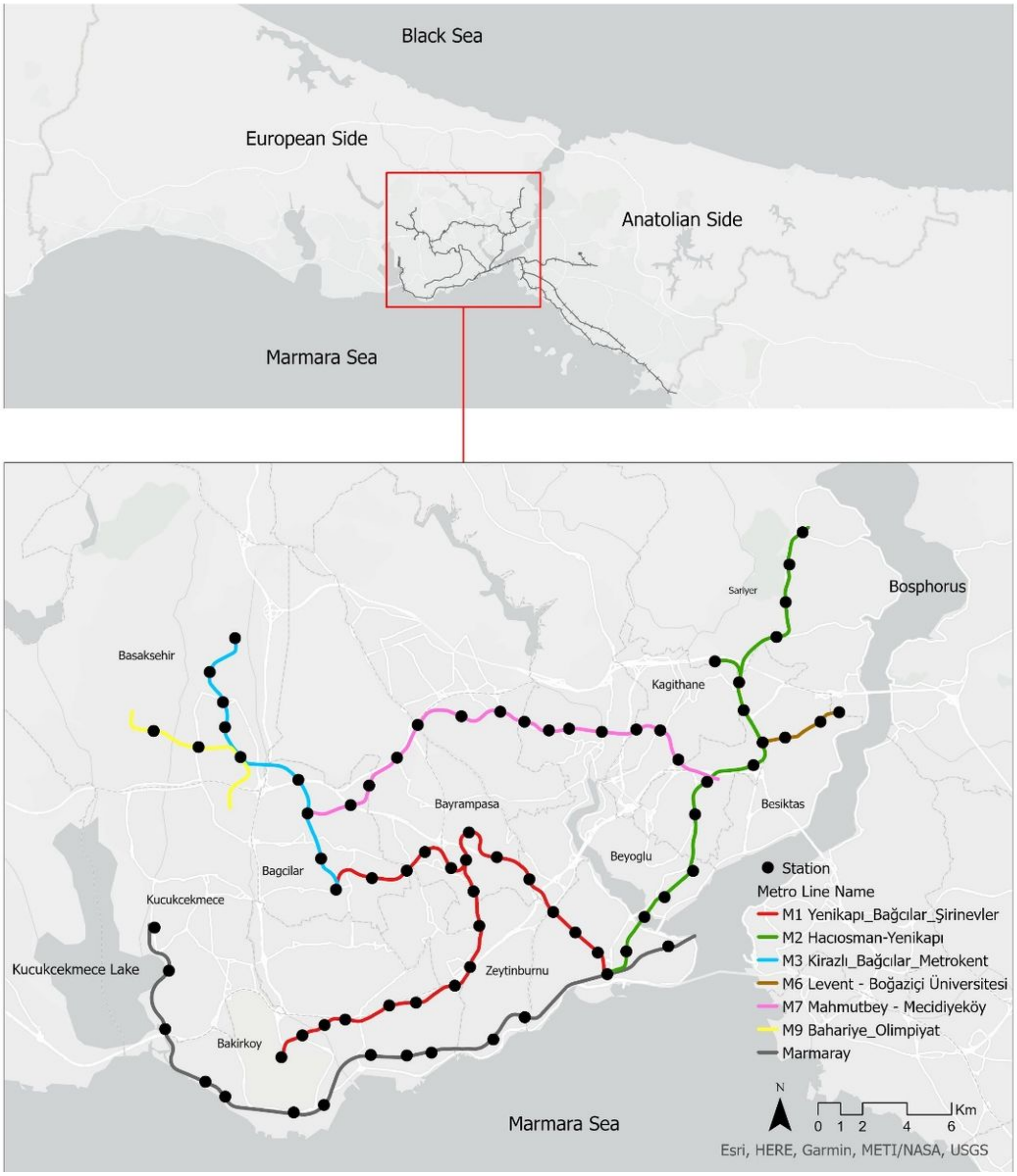


Figure 1

Study Area

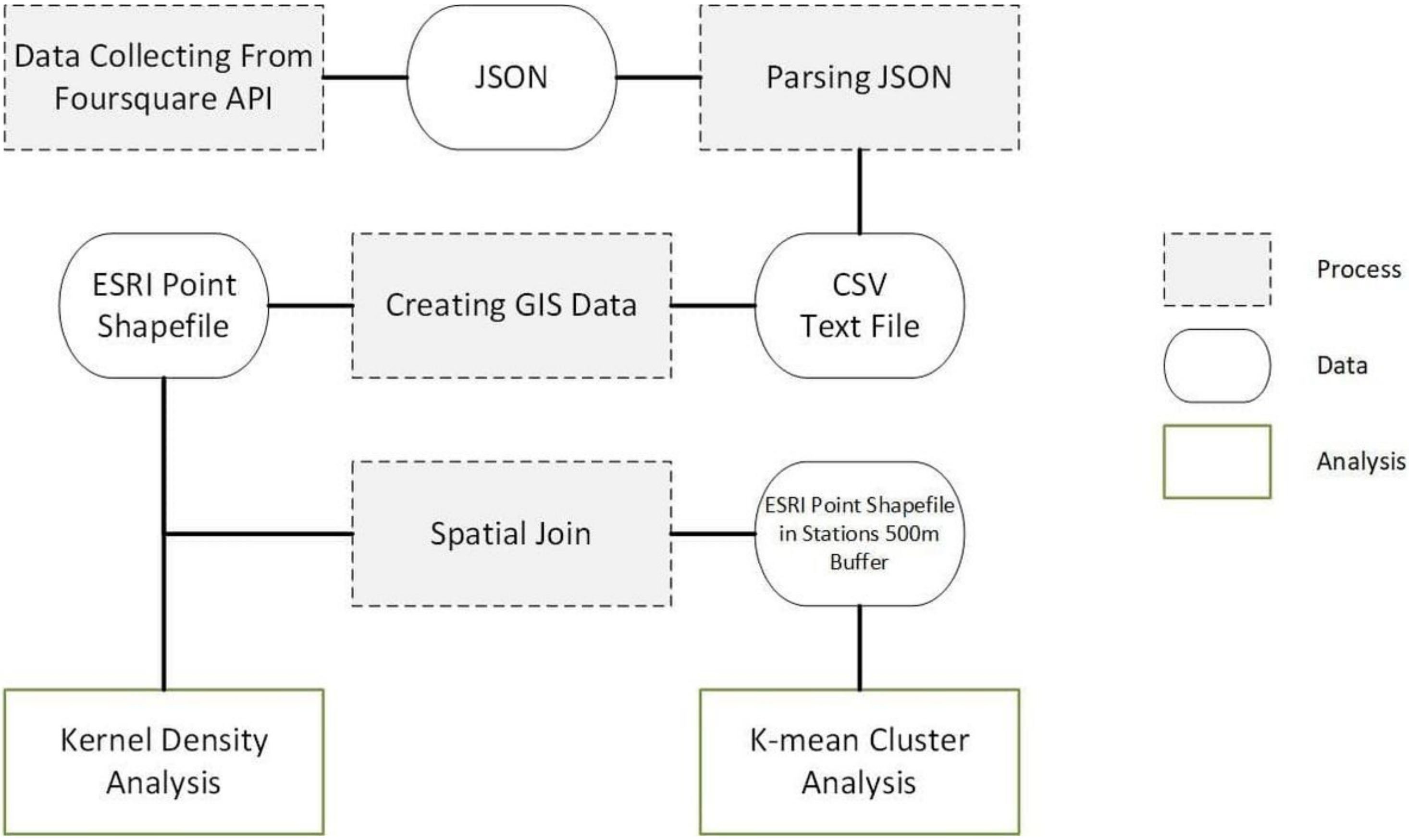


Figure 2

Methodology flowchart

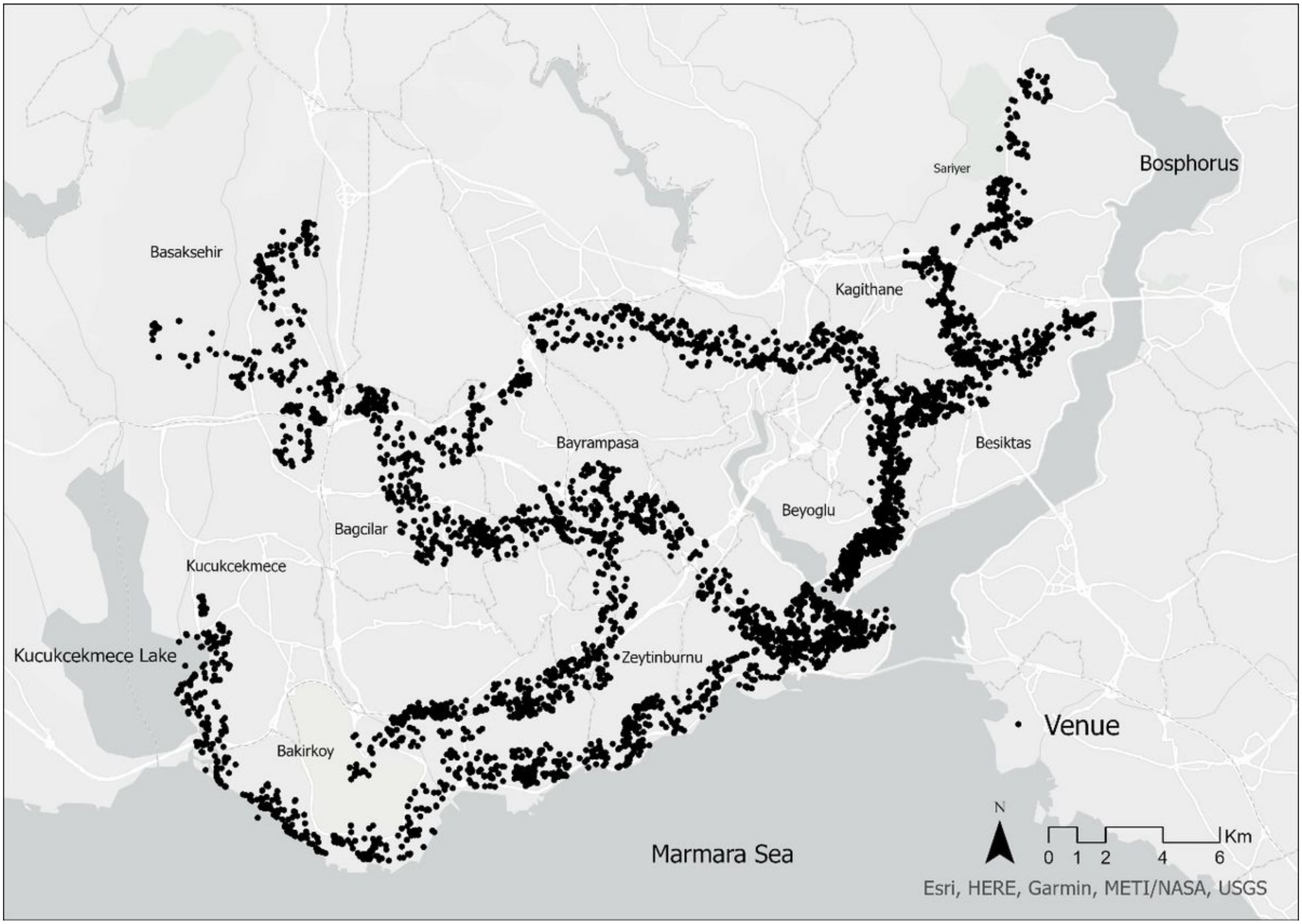


Figure 3

venues extracted from foursquare

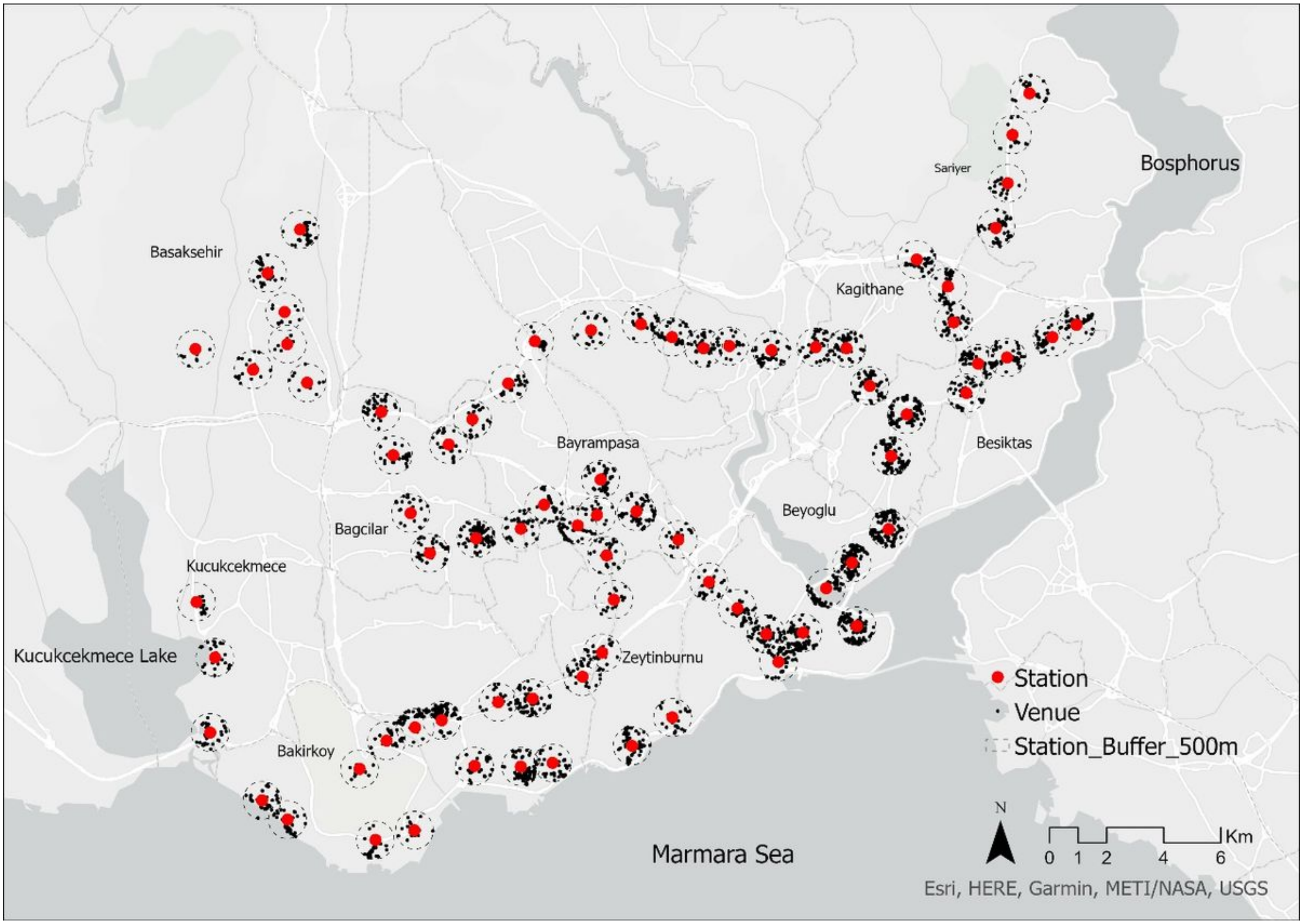


Figure 4

Venues extracted from Foursquare around metro stations

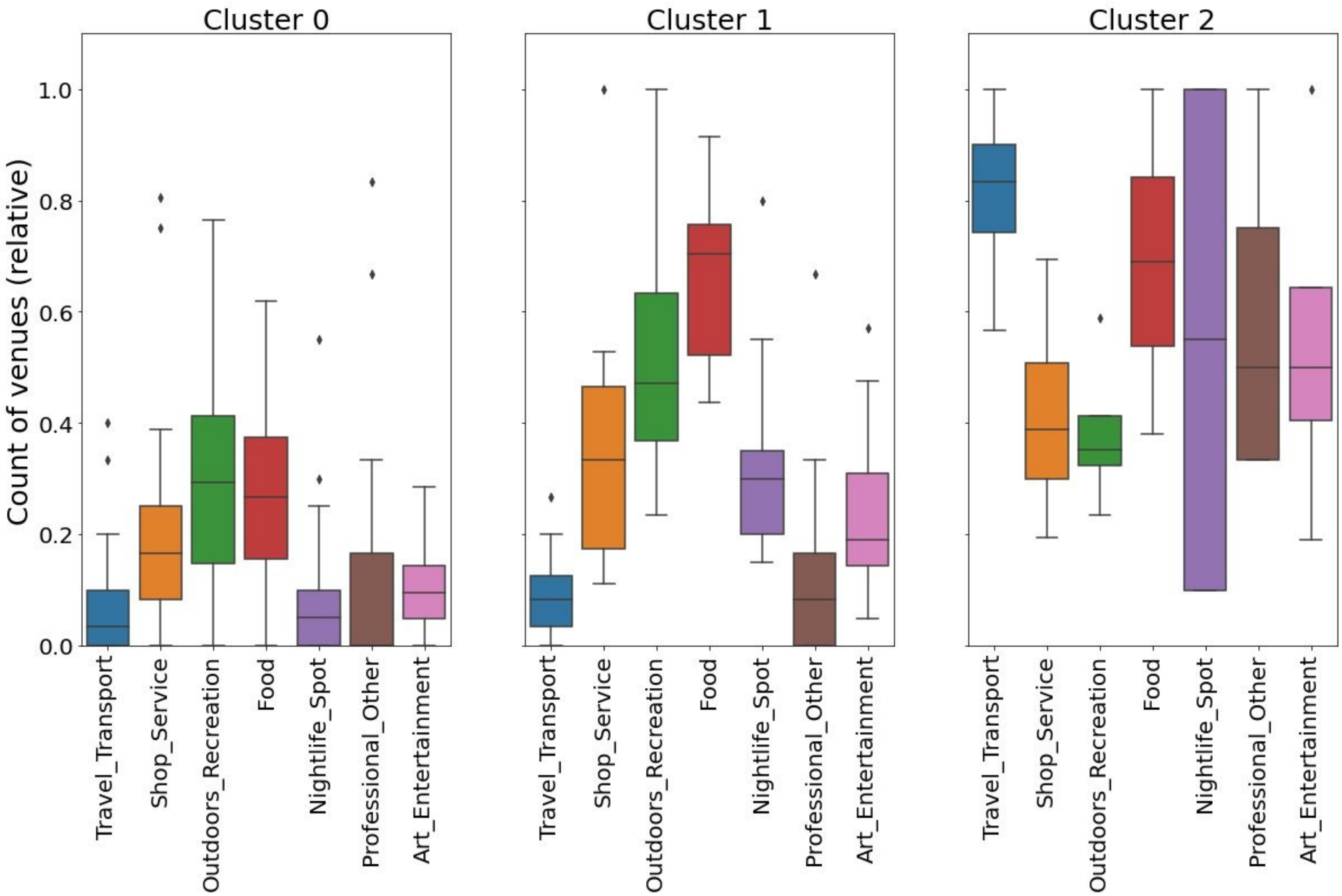
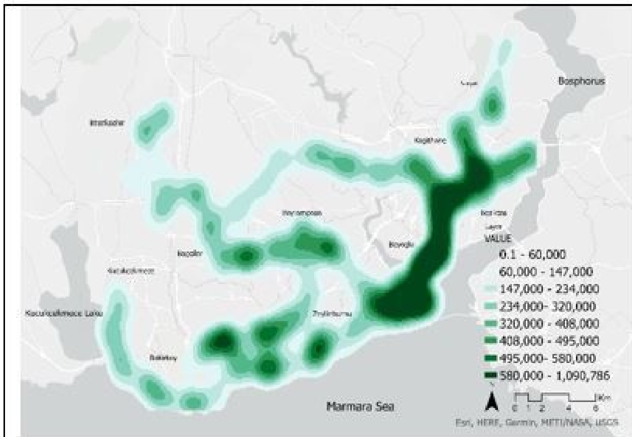
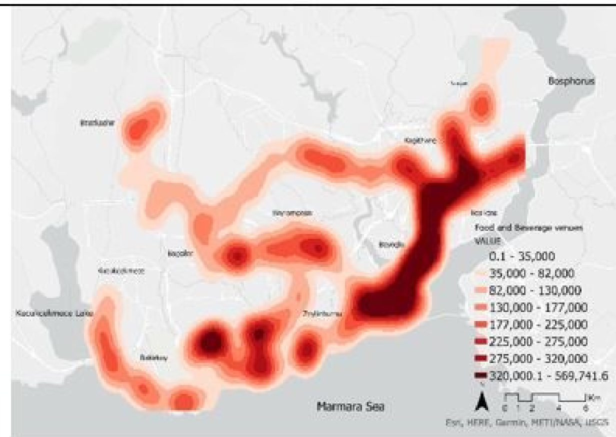


Figure 5

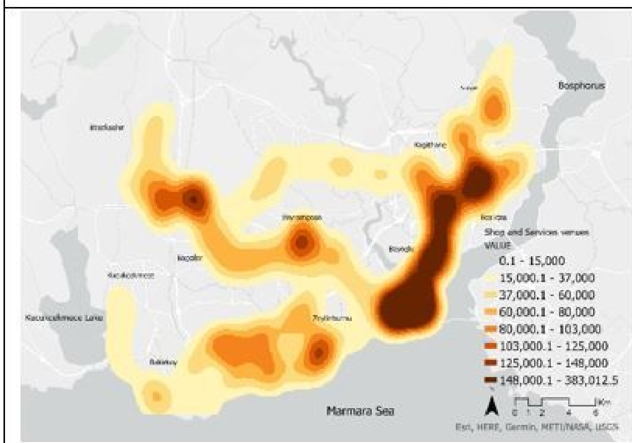
Statistics of each cluster in the boxplot



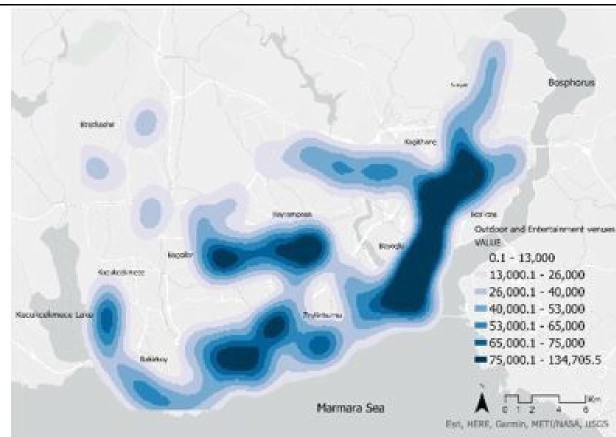
All venue



Food and Beverage venues



Shop and Services venues



Outdoor and Entertainment venues

Figure 6

Kernel Density Analysis

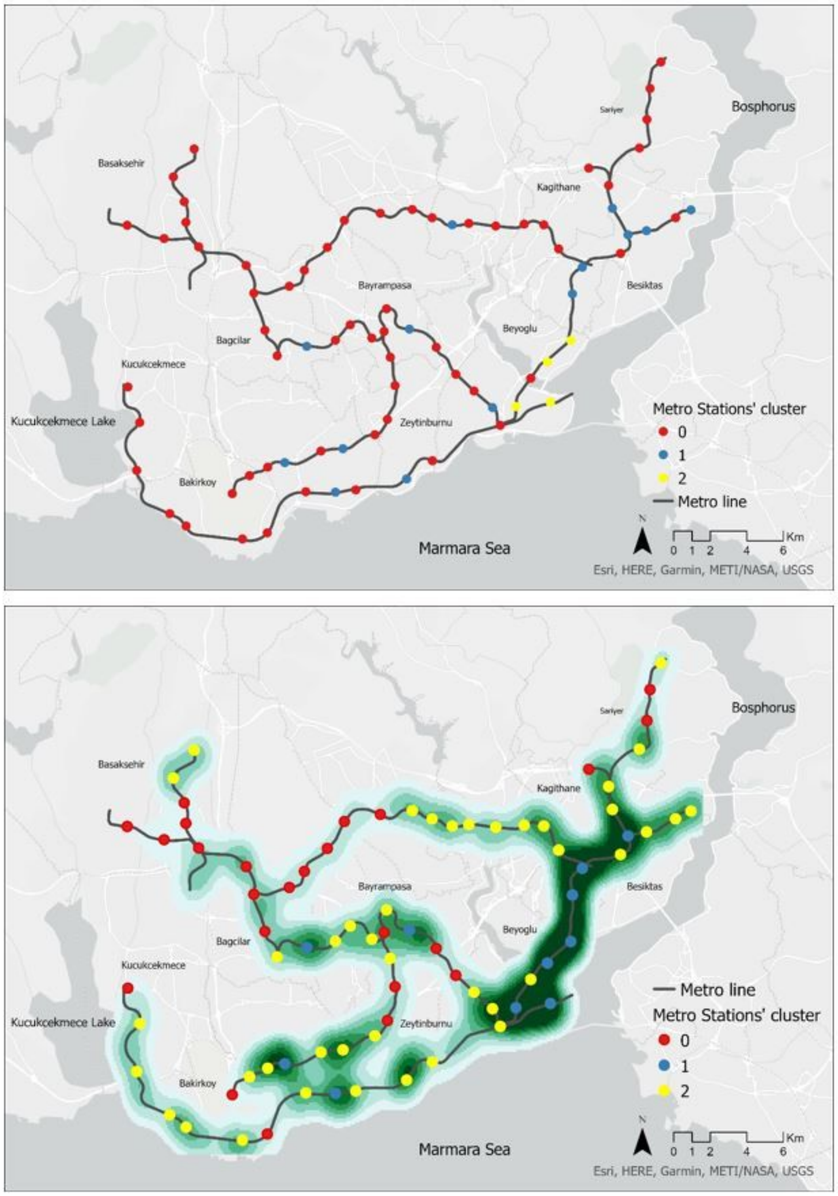


Figure 7

Cluster analysis of metro stations