ITU A|Z • Vol 15 No 1 • March 2018 • 5-21

Web scraping and mapping urban data to support urban design decisions

Elif ENSARİ[1], Bilge KOBAŞ[2]
[1] elif.ensari@gmail.com • Architectural Design Computing Program, Istanbul Technical University, Istanbul, Turkey
[2] bilge@super-eight.com • Bits'n Bricks Analytical Research Center, Istanbul, Turkey

Received: October 2017 • Final Acceptance: November 2017

Abstract

Cities generate data in increasing speed, volume and variety which is more easily accessed and processed by the advance of technology every day. Consequently, the potential for this data to feedback into the city to improve living conditions and efficiency of utilizing resources grows. Departing from this potential, this paper presents a study that proposes methods to collect and visualize urban data with the aim of supporting urban design decisions. We employed web scraping techniques to collect a variety of publicly available data within the Kadıköy municipal boundaries of Istanbul and utilized a visual programming software to map and visualize this information. Through this method and superposition of our resulting maps, we visually communicate urban conditions including demographic and economic trends based on online real estate listings as well as spatial distribution and accessibility of public and commercial resources. We propose that this method and resulting visualizations present valuable potential in supporting urban design decision-making processes.

Keywords
Geospatial data, Urban data, Urban design decision support, Web scraping.

doi: 10.5505/itujfa.2018.40360

1. Introduction

Data turns out to be more relevant for urban design as it becomes more ubiquitously available and accessible through the advance of technology and as it is more widely incorporated in the decision-making processes regarding the urban environment. Recent years have seen the proliferation of big data as well as research concerned with its integration into the analysis, management and design of cities. Numerous academic, civic and commercial research centers have been founded dedicated to studying the urban environment through information generated by the cities (Batty, 2013). While the scale, variety and production speed of data (3Vs) increase every day, its availability, accessibility and communication to designers, decision makers and to general public remain a critical issue. This is due to the range of skills required to access and process the data in order to make it relevant and meaningful.

The aim of our study is to present an urban data collection and visualization methodology which can help urban designers produce maps to communicate various urban patterns to stake holders and decision makers in order to support urban design decisions, as well as to communities to better inform participatory processes. We also present a set of maps produced following this workflow, along with some basic geospatial analyses we have performed to create them. The questions we take on to answer through the study presented in this paper are as follows.

- How can sources of publicly available data with geolocation information be relevant and useful to urban designers?
- How can this data be gathered and visualized?
- What kind of knowledge can be generated through the visualization of this data that may be presentable to decision makers even before any geostatistical analysis is carried out?

The study is unique in that we present a methodology to compile a rich dataset which we propose to be relevant for urban design decision making processes, consisting of information drawn from multiple sources rather than studying a single source of data and its relevance for a specific urban issue. We also present visualization of each data set with some basic spatial analysis. Through this, we aim to aid quick and easy visual analysis and better communication of this data to designers and decision makers. Furthermore, Kadıköy region of Istanbul has never been subject to such a comprehensive study in terms of the variety of urban data and data sources utilized as well as the techniques adopted for mapping and visualization. The data collection was done utilizing Python scripting and all mapping, visualization and visual analysis processes were done using the visual programing environment Grasshopper for Rhino3d.

Our research has revealed various urban patterns that became apparent through the maps we have created utilizing the data collected from multiple geolocated urban data sources. Two most significant cases we have observed to have potential in contributing in the urban design decision dialogues are the exposition of public and commercial resource distribution, and demographic and economic trends read through the various maps we present in the following sections.

We aim to develop this study by integrating more comprehensive analyses to our workflow in future research.

2. Literature review

While cities around the world are growing in size and population every day, digital devices and systems are becoming more and more integral to their functioning in parallel with technological advancements. This results in the generation of massive amounts of data in the urban environment which have already been subject to a large body of research motivated with the belief that cities can learn from this data and become 'smarter' in order to provide better living conditions for their residents while utilizing resources more efficiently (Batty et al., 2012; Townsend, 2013). Mobility patterns of urban dwellers drawn from digital public transportation ticketing systems, cellular phones or cameras detecting traffic patterns, GIS data compiled and made available to public by governments, various use and activity data shared through personal mobile devices and publicly available social media content are some of many examples of such data.

A part of these studies in urban data focuses on gathering, analyzing and visualizing data which is inherently linked to geolocation information, namely geospatial big data (Lee & Kang, 2015), making it possible to map and observe geospatial patterns of urban phenomenon. While such studies often bring together geographers and computer scientists based on interest and technical skills required; urban planners, designers and architects also see value in this data which can provide precious insight on the functioning of cities and support urban decisions in order to improve the quality of the urban environment (Glaeser, Duke-Kominers, Luca, & Nalk, 2015).

One of the most popular subjects of study of dynamic urban data is real estate rental and sales prices, mapping and geographical analysis of which allow for understanding location driven contributors to property values (Cohen & Coughlin, 2008; Geoghegan, Wainger, & Bockstael, 1997; Waddell, Berry, & Hoch, 1993). More recently, studies utilize housing price information from public real-estate websites (Bency, Rallapalli, Ganti, & Srivatsa, 2017; Boeing & Waddell, 2016), facilitating the use of very large and up to date datasets, which allows for the development of more accurate housing price prediction models. One such study finds correlations of housing prices with geo-tagged social media content (Li, Ye, Lee, Gong, & Qin, 2017), which is another type of data that has been subject to urban analysis research as an indicator of urban activity (Jenkins, Croitoru, Crooks, & Stefanidis, 2016; Schreck & Keim, 2013). Walking, running and other forms of exercise data tracked and shared by urban residents have also been studied to better understand qualities of preferred locations and routes for recreational activity in the urban environment (Balaban & Tuncer, 2016; Clarke & Steele, 2011).

There are various methods and techniques which allow for management of big data, yet means of its collection and visualization is the main subject of our interest in this research. As a method to draw data from public web sites, web scraping which we utilize in our study involves automated means to query and download large sets of geospatial data. Boeing and Waddell (2016) utilize this technique to analyze real-estate listings in the US and are able to compare eleven million listings all over the country with government declared fair market rents in 58 metropolitan areas. APIs (application programming interfaces) provided by many online platforms such as Google, Instagram, Facebook, Flickr, Foursquare or Twitter make it possible to query and download user generated content, which becomes interesting for urban researchers when the data is accessed together with geolocation information. Among several recent studies that utilize social network data to better understand urban patterns, Cranshaw and colleagues (2012) detect areas of concentrated urban activity ("Livehoods") using Foursquare check-ins and Jiang, Alves, Rodrigues, Ferreira, & Pereira (2015) estimate land use based on Yahoo's point of interest (POI) data with a higher accuracy than with traditional methods. Chen (2014) provides a good overview of data collecting methods from social media platforms including web scraping and the use of APIs that were also utilized in our study.

All geospatial data is already big data (Lee & Kang, 2015), and velocity, variety and volume are the 3Vs considered to define big data. While collecting, managing and analyzing such rapidly generated large scales of data already pose a significant challenge, visualizing and communicating it is utmost critical in deriving meaning and value from geospatial data in order for cities to learn from it. There are several tools from geographical information systems (GIS), building information models to virtual simulation environments that make available various visualization techniques (Pettit et al., 2012) for geospatial data.

A project that demonstrates the power of visualization of geospatial data, "Million Dollar Blocks" (Kurgan, 2013) maps the home addresses of criminals as they are admitted into prison as well as the length of their stay in five of US's largest cities and reveals

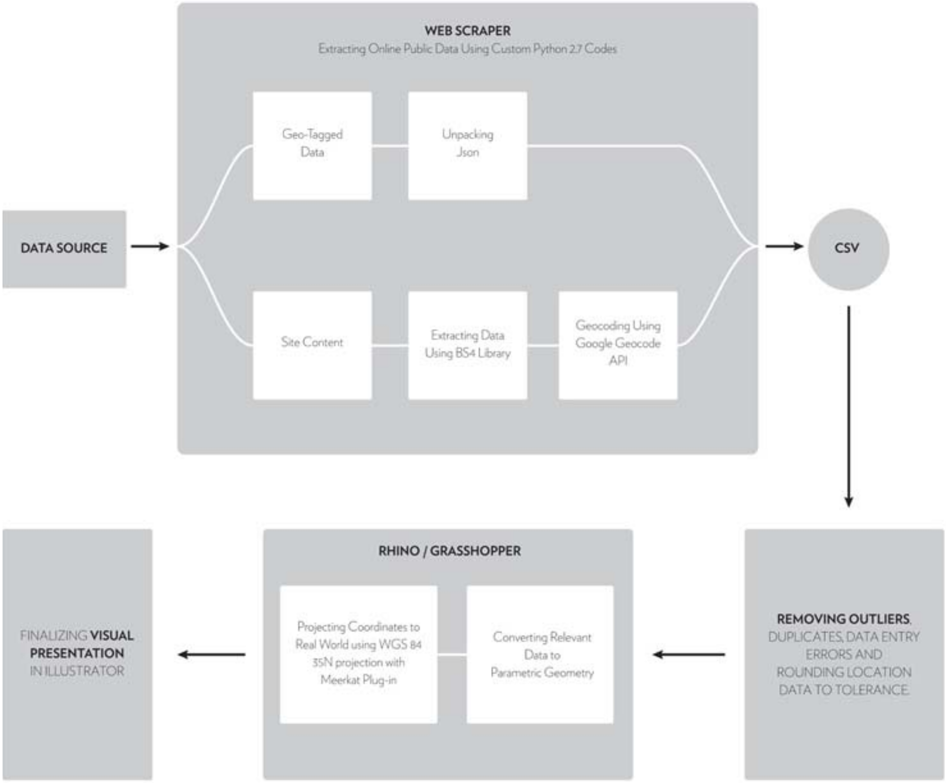

Figure 1. *The workflow.*

a remarkable concentration of these addresses around a few city blocks. One of the project's most striking visuals maps the money spent by the state on each criminal to their home address, displaying 4.4 million dollars spent on criminals coming from within just four blocks in Brownsville, Brooklyn. A study conducted in a newly activating area of Amsterdam, The Knowledge Mile (Niederer, Colombo, Mauri, & Azzi, 2015), presents geographical maps of companies and their online hyperlinks, most shared images and most shared locations using data gathered from online platforms Google Search, Panoramio, Instagram and Foursquare. Making apparent the online presence and character of the urban axis in the focus of the research, these maps were later shared with local stakeholders in participatory design sessions and reportedly initiated conversation and further field work.

Having various similarities with the research presented above, our study presents multiple layers of geographically linked urban data drawn from dynamic sources of public web content, mapped to reveal patterns of urban activity and distribution of commercial and public facilities. We propose that these visualizations created through basic spatial analysis and mapping techniques allow for human visual assessment of a rich set of urban data, which is deemed as valuable as automated analytical processes (Schreck & Keim, 2013).

3. Methodology

Our process involves the following steps: drawing of data from online sources, organizing the data, performing mathematical operations to calculate some additional values on the database, mapping the data based on geo location information, performing basic geometrical analysis and visualization. Our workflow is presented in Figure 1.

We utilized multiple sources of publicly available data from within the Kadıköy municipal boundaries of Istanbul, to draw already geo-tagged data or data with address information that we converted to latitude and longitude values through Google's Geocode API. While mapping and visualizing the raw data we collected, we applied different levels of intervention; from simply positioning the circles at geographical locations and scaling them based on their values within the domain of the data set or connecting the geolocation

Figure 2. *Unit rental prices (rent price per sqm).*

points to create routes and superposing them, to performing surface to surface intersections to calculate areas of influence or accessibility of amenities automated through custom codes.

Our data comprises real estate rental and sales prices, property square meters, property bedroom numbers, ages of buildings, uses of buildings, social media activity, public park boundaries, commercial facilities, public transportation and other public facilities.

For the data collection phase, we use Python programming language to send requests composed by the Postman to the source web site's servers and generally download the data in Json format. We then convert the data to csv (comma separated value) format and import it to Grasshopper visual programming environment, an add-on for Rhino3d, using Meerkat plugin. Here we process the values within the database or run simple spatial analysis after converting them to geometrical entities. All the steps carried out in the Grasshopper environment besides the final graphic touch-ups are automated so that they are easily repeatable for different locations in future studies.

Below we present the details of our methods grouped based on the sources of data and the themes of the maps that were created.

3.1. Housing and real estate

We utilized one of the most widely used real estate listings websites in Turkey with 42.6 million active members visiting the portal 263 million times per month that allows for posting and managing of apartment rental and sales listings by both the owner of the property as well as real estate agents. The website has a map interface through which we were able to draw geolocation, size (in sqms) number of rooms and comments of the listing owner of each apartment in the Json format. Then we sent a request to the page of each listing to draw further text based information including the type of heating, additional amenities and whether the listing was posted by the owner or the real estate agent. The text data was parsed using the Beautiful Soup Library for Python and converted into csv format with Csv Library for Python. The csv files were then imported to Grasshopper and processed to calculate the rental and sales prices per sqm and create maps (Figures 2 & 3). In these maps, the radius of each circle positioned at the geolocation of the

Figure 3. *Sales prices per sqm.*

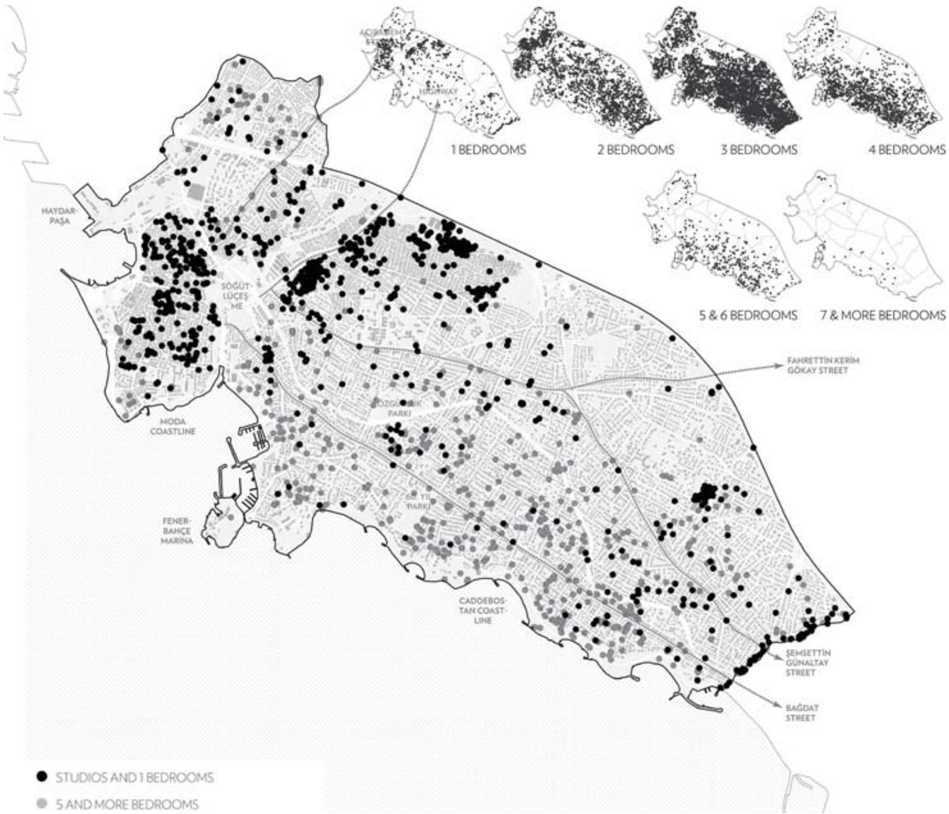

Figure 4. *Number of bedrooms for each house listing (rental and for sale).*

listing is scaled proportionately with its price divided by sqms, after the outliers were excluded from the dataset. The number of rooms per apartment and building age data were combined from both rental and sales listings and mapped with colors (Figures 4 & 5).

Payback values were calculated per 100m by 100m area based on the average apartment rental and sales prices

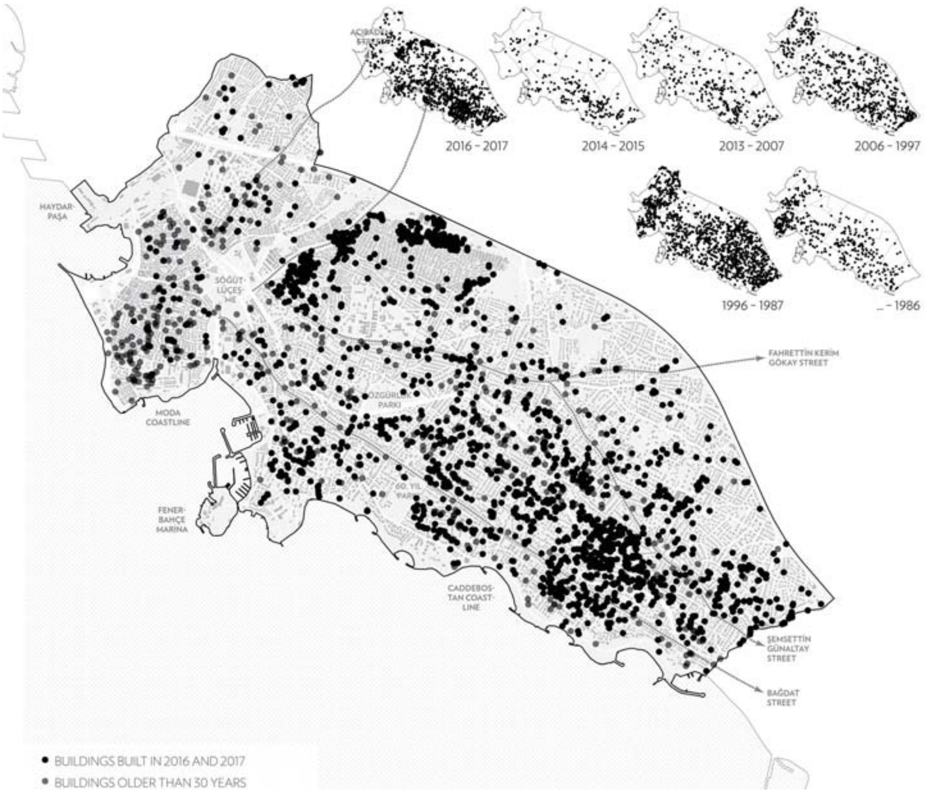

Figure 5. *Ages of the buildings of the listings (rental and for sale).*

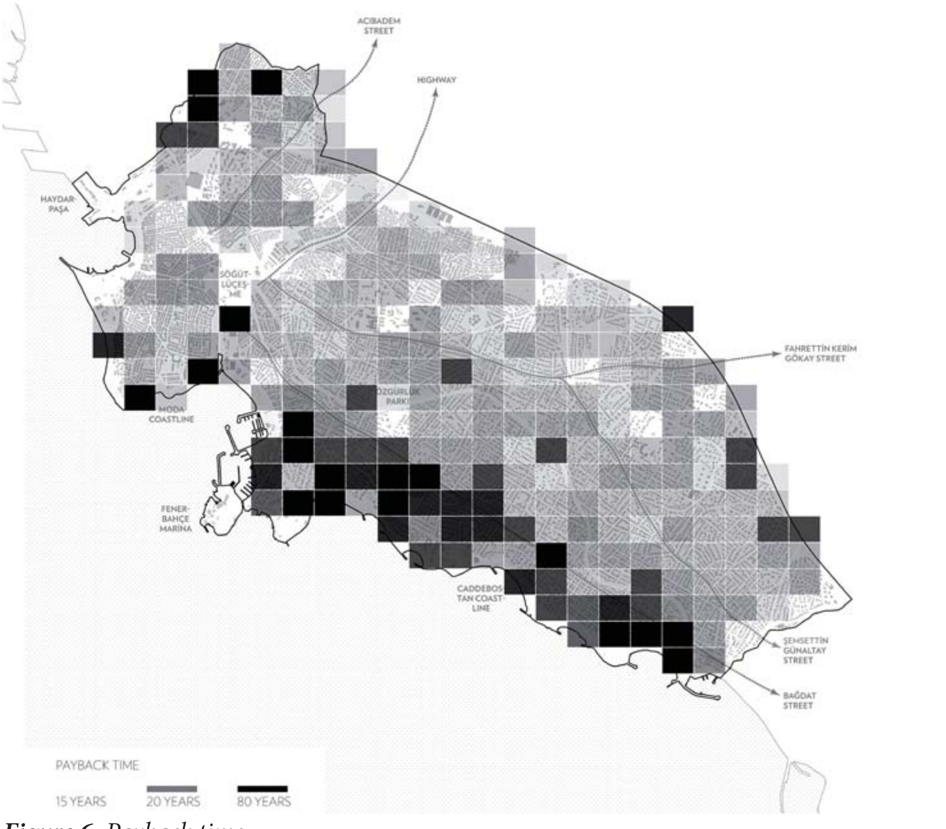

Figure 6. *Payback time.*

per sqm available in each grid square (Figure 6).

A GIS file created by Istanbul Metropolitan Municipality was used to map the residential/non-residential building use information, with the building footprints colored using a gradient based on the proportion of the number of non-residential units to the residential ones for each building (Figure 7).

Figure 7. *Residential and non-residential uses in buildings.*

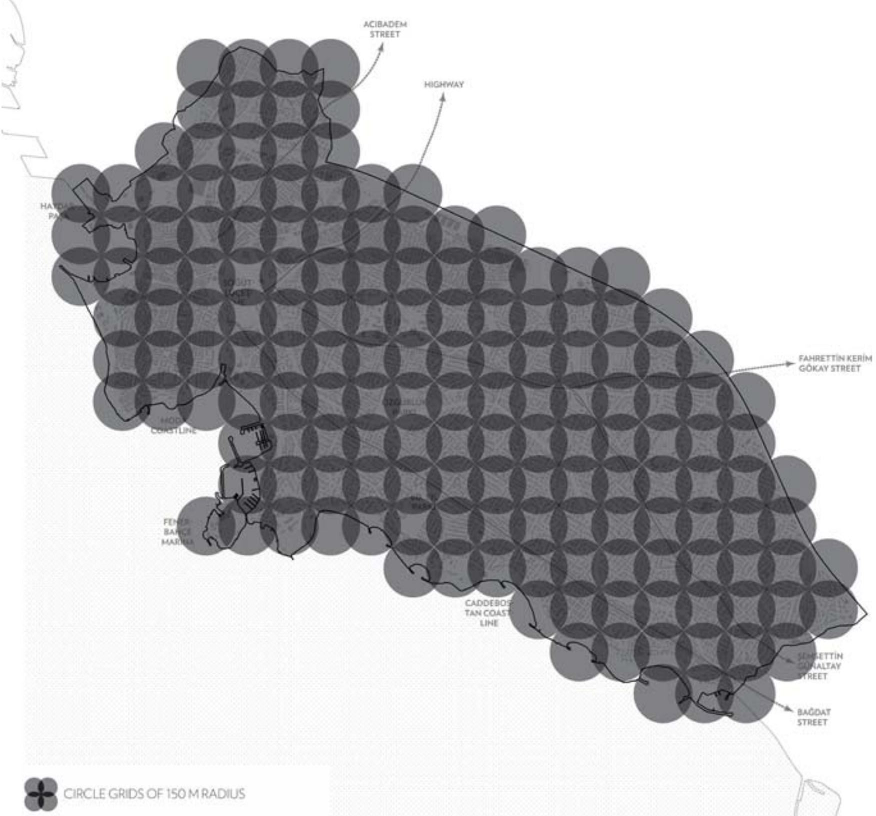

Figure 8. *Circle grid system for Instagram queries.*

3.2. Social media

Instagram was the source of social media data we utilized for this study as it is known to be favored over Twitter and other forms of social media in Turkey, with 16.34 million users by 2016 (Statista.com, 2016). Python code and Instagram's own API was used to query 100 posts at a time, from within 150 m radii of 159 points arranged to cover all

Figure 9. *Cafes, Instagram posts and their like numbers.*

the area within the municipal boundaries of Kadıköy (Figure 8). Besides the geolocation data, number of likes and the text accompanying each image was drawn. The duplicate posts downloaded due to inevitable overlaps of circles were cleaned up. Posts that were liked over 1200 times were counted as 1200 due to their low frequency and posts within a distance of 5m's to each other were combined by adding up their number of likes for the sake of a more legible representation. A map was created with the scale of each circle representing the like numbers per post point (Figure 9).

The text data drawn from the Instagram posts were analyzed using a custom definition separating each word using Python split function defining space sign " " as a delimiter and hashtags were appended to the data of each post in the csv file. Then the frequency of each hashtag was calculated per 100m by 100m grid and mapped (Figure 10).

3.3. Public and commercial amenities

Using the Google Places API, we drew the geolocations and ratings of Cafés and Restaurants which were marked on Google Maps. Similarly with the method we drew Instagram posts, we queried 159 points with 150m radii and recorded 60 places at a time. Then we cleaned up the duplicates and created a simple point-location map.

One of our resources was a widely used food delivery portal, Yemeksepeti.com, with over 5.2 million users by 2016 (2017), that directs orders from users to member restaurants in Turkey. Each item on a restaurant's menu is displayed with its price, however, the restaurant's address is not provided on the website. First, we queried the geolocation information of each restaurant using Google's Geocode API. Then the prices of the most commonly sold items which were a 330ml can of Coke and 250ml cup of Ayran (a yogurt based Turkish beverage) were queried. We also drew the hours of operation for each restaurant. The website does not provide the menu items in a database format therefore a custom Python web scraper script was developed which searches for all restaurants delivering to the area and then follows each of their links to access their menus. Beautiful Soup Library was used to convert the HTML formatted site into a formatted spreadsheet, the prices were converted to the scales of circle radii distributed

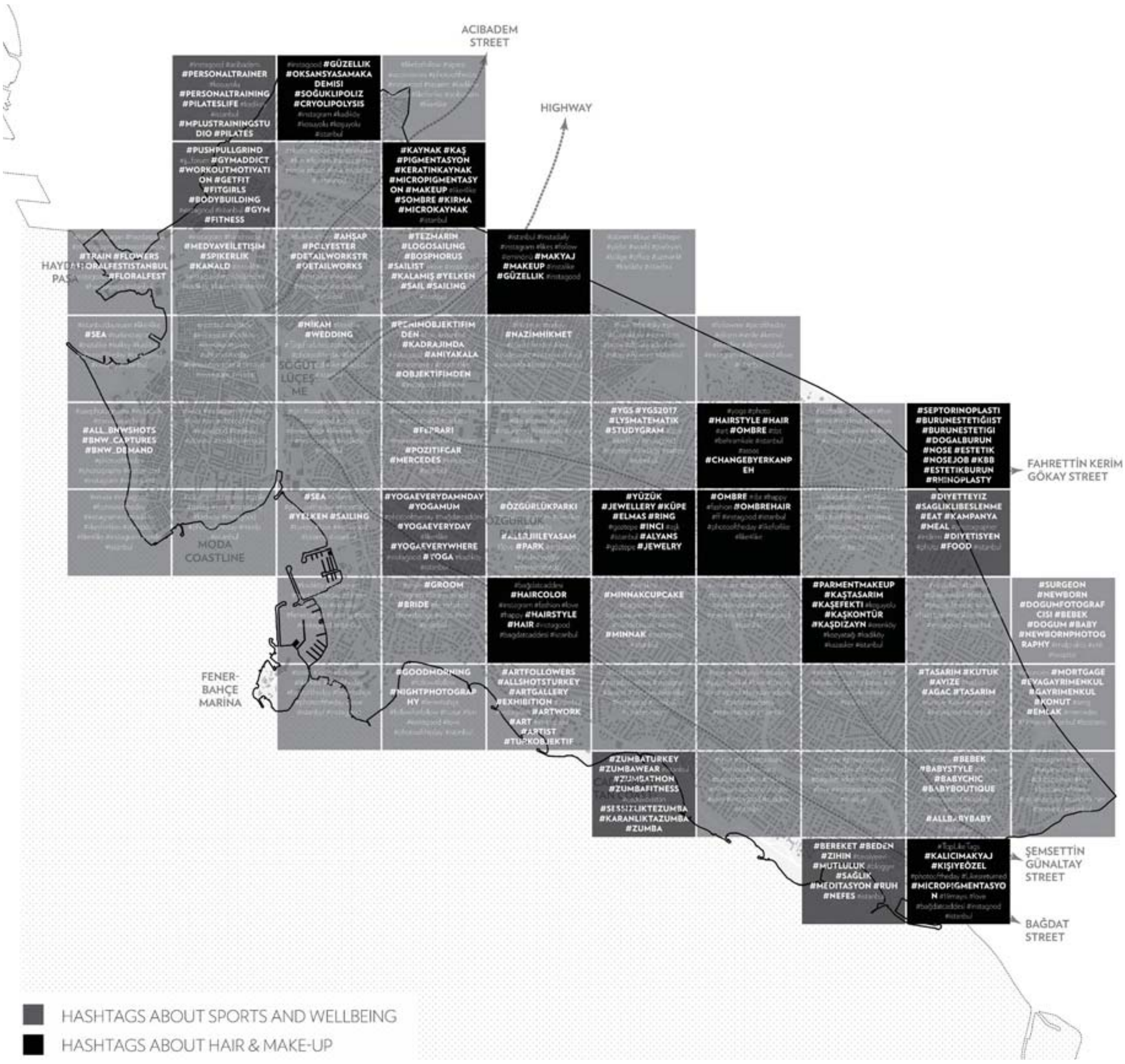

Figure 10. *Hashtag map.*

on the map based on the geolocation of each restaurant and the number of operating hours were converted to color codes to be visualized on the map (Figures 11&12).

3.4. Transportation and access

We drew the geolocation data of stops for buses and metrobuses (a rapid bus transit line using dedicated bus lanes for much of the routes) from IETT's (the general transportation department of the City of Istanbul) website as well as the number of times a bus or metrobus passes through each stop to create routes on a map (Figure 13). We used their frequency data to define the thickness of the route lines on our map.

Then for each building centroid, we computed the number of public transport seats accessible within 400m, 800m and 1200m radii. For this, we calculated the number of stops within these circles first, and then multiplied the seated passenger capacity of each vehicle with its frequency throughout the day. Without drawing the route lines, we did a similar calculation for ferries and subway lines, and added up all the seated passenger capacities within the mentioned radii distance from each building centroid. We illustrated these numbers through a gradient of colors for each building, representing a scale of accessibility to public transportation. (Figure 14)

Another map we created aims to visualize the accessibility of recreational green spaces from buildings. For this, we utilized the geographically located outlines of public parks drawn from the GIS files and made surface intersections of them with the circles of 400m, 800m and 1200m radii from

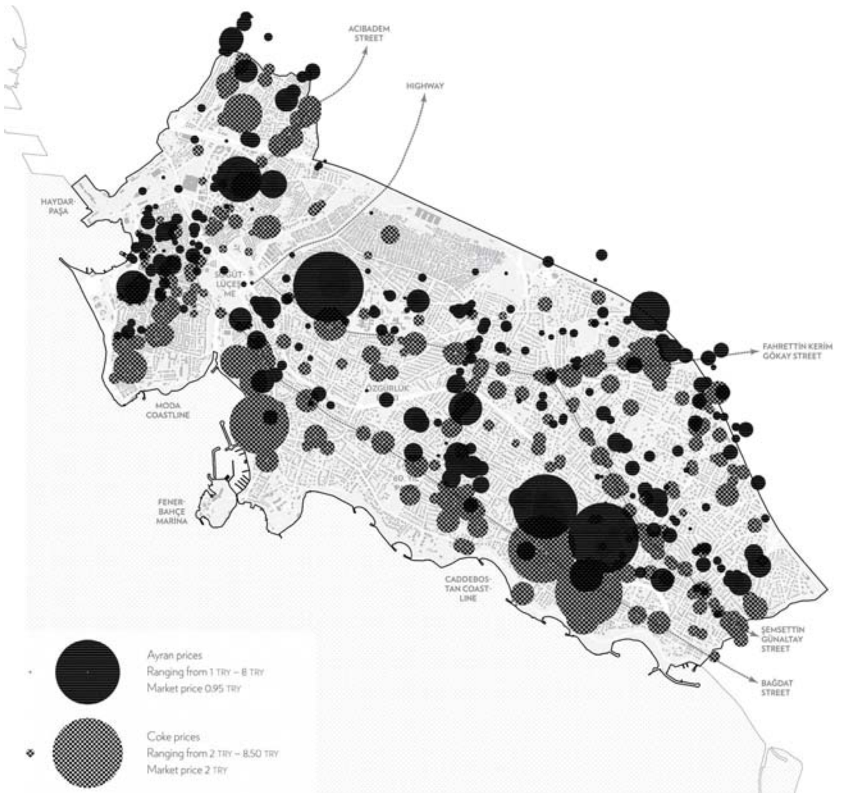

Figure 11. *Soft drink prices.*

Figure 12. *Operating hours of the restaurants.*

each building centroid. The areas of intersecting regions gave us a domain of accessible green space sqms that we mapped as a gradient of colors for each building polygon (Figure 15).

We drew the exercise routes of Kadıköy residents from an activity tracking social network that allows for sharing of walking, running, biking and swimming activity data. Superposing these routes revealed

Figure 13. *Bus and metrobus lines.*

Figure 14. *Accessibility to public transportation seats.*

the most popular streets and in one case a swimming route, as well as the approximately estimated geographic locations of the buildings, residents of which use these routes most frequently (Figure 16).

4. Observations and urban design decision support

We propose that the set of maps produced by our automated workflow are powerful in demonstrating how urban

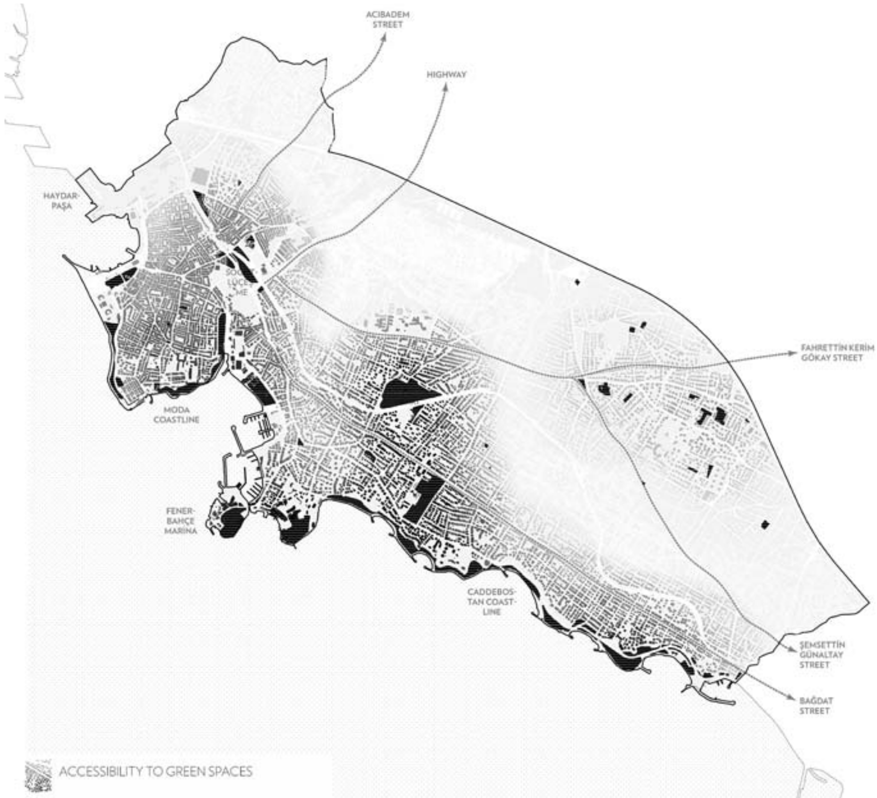

Figure 15. *Accessibility to green spaces.*

Figure 16. *Exercise routes.*

data can be basically organized and visualized to communicate meaningful observations and aid urban design decisions.

Although the quantitative results we aim to generate in future studies through geospatial analysis of this data will advance this research significantly, we believe that the current stage of the study we present here is as useful in revealing the relevance of geospatial information for urban design. Some of our observations at this stage of our research are presented below. We also present ways in which the visual information revealed in these maps can support urban planning and design decisions following each of our observations.

There is a concentration of apartments for sale and higher sales prices along the coast and a concentration of rentals of 1+0 and 1+1 units on the west of Söğütlü Çeşme (Figures 2, 3 & 4). This is informative about the demographics of residents around these neighborhoods and can guide the municipality in deciding on the kind of investments they should make, may they be introducing co-working spaces, libraries, professional education facilities, kindergartens or retirement homes. Our Instagram hashtags and soft drink prices maps (Figures 10 & 11) similarly contribute in revealing the demographics of the local residents and visitors, as well as providing a glimpse of their consumer trends.

Urban transformation which is encouraged by the policy to renew non-earthquake resistant housing stock of Istanbul is made apparent in our maps displaying the ages of buildings (Figure 5), showing the geographic distribution of recent construction. The rapid transformation within the Caddebostan area in the last couple of years is revealed which may be obvious to the neighborhood residents but no known maps exist with updated information available to public. The local authorities can utilize this information not only for understanding areas where transformation is slow, but also where it is rapid, and they can work with urban designers to take precautions against disadvantageous consequences of such a high density of construction sites and the rapid change of the built environment for the local residents.

Together with the cafes & Instagram likes map (Figure 9), our residential-nonresidential buildings map reveal (Figure 7) the main commercial axes and the mixed-use pattern in Kadıköy which is known to be an indicator influencing the walkability in a neighborhood. The operating hours of restaurants (Figure 12) influence the liveliness and feeling of safety on a street, contributing to the "Eyes on the street," a concept first proposed by Jane Jacobs (1961). The geographical distribution of commercial amenities and social media activity can inform urban decisions of zoning and investment in new facilities or services by revealing if the existing facilities are sufficient or attractive to urban dwellers and how commercially active a street is throughout the day. Areas that are lacking in such amenities can be prioritized for making changes in planning to encourage a more vibrant street life and commercial activity on their streets.

Our public transportation and access maps (Figures 13 & 14) visually present one of the most common indicators of walkability for urban neighborhoods. Although the distance from each building was not measured through the street network but simply calculated based on a radius, the map showing access to public transportation seats provides an idea of how democratically distributed the transportation facilities are within the municipal boundaries of Kadıköy. Access to greenspace that we created a map for (Figure 15) is a similarly significant contributor to walkability and quality of life. Our exercise routes map (Figure 16) gives an idea of which residential locations have better access to recreational activities within the city. Besides distance, we argue that this map may reveal additional contributors to the exercise behavior in the city such as the pleasantness, comfort and safety of these routes, as well as their accessibility from different neighborhoods.

Municipalities and central authorities can manage the allocation of their resources based on information regarding the distribution of public services and facilities easily observable through our maps demonstrating access to public transport, green areas and routes used for sports activity. Since residential neighborhoods less fortunate in terms of access to public transportation can quickly be identified in these maps, they can encourage local residents to demand an increase of routes and more frequent services from transportation authorities. Access to green areas and the shoreline which is obviously attractive for exercisers can be enhanced by new bike lanes and improved street conditions for better walkability especially in areas that link the northern parts of Kadıköy to these areas.

We present these observations and suggestions for urban designers through the case study of Kadıköy, however, we aim this study and these maps to constitute a model for revealing and communicating information that can guide urban decisions in improving the street life, commercial activity, local development and life quality of urban residents through better allocation and more democratic geographic distribution of resources. Without doubt, each urban area from each country is unique in many aspects that also contribute to the urban qualities that we propose can be improved through our observations, thus all data should be mapped and interpreted with various such differences in mind to be relevant in supporting urban decisions.

5. Discussion

We do not present any geostatistical analysis of urban data in this paper, but rather propose this research as a preliminary step to be followed by rigorous spatial analyses in studies to follow. Nevertheless, we propose the methodology we present for urban data collection and visualization as a guide for designers to better understand and utilize means to gather and process data relevant for urban design decision support.

We acknowledge that there may be possible errors in information provided by the listing owners regarding the real-estate data we utilize. Voluntarily generated data may also cause inaccuracies due to the free data-entry format, as in the case for social media entries. The social media users tend to create their own use of language, abbreviations, and personal methods of using hashtags (Schreck & Keim, 2013) or simply make misspellings. Furthermore, Turkish is an agglutinative language and multiple suffixes make it unfavorable for text based analytics.

Additionally, the sources we have used for this research are limited: There are several websites and applications dedicated for similar purposes, and at this stage of the study, the datasets were created using the most widely used website and/or application in their respective category. While this approach gave us a strong initial start, additional sources may be added to the datasets in further research. In such a case, a mutual database format should be created as the type of data retrieved from different sources will also begin to differ.

Final discussion marks include copyright issues; as almost all of the websites used in these studies limit reproduction and re-distribution of the data for commercial use in their legal user agreements. Boeing and Waddell (2016) note a case where a federal US court decided that web scraping publicly available data did not constitute a copyright violation. Even though non-commercial research does not fall under legal restrictions, only a handful of the websites and applications make the data collecting process openly available to developers and researches. This means that the sources without developer access may easily limit or restrict our access in the future by making changes to their web structures.

6. Conclusion

Nowadays, a significant portion of public and commercial activity takes place on online social networks. We swim in an enormous pool of data streaming all around us. The tweets, selfies, check-ins, likes, ignores, pokes, snaps, listings, stars, user comments constantly keep updating and adding to this pool. Mostly consisting of unorganized, unfiltered and uncategorized raw data; this pool is the ultimate collective subconscious of our society. It has the information on what people like doing where, when, how; how they travel; how they spend their time; how they speak to each other; what their share of economy is; how they take what was designed and how they are transformed by that. This is a valuable source of information for anyone interested in exploring the geospatial past, current time and future in great detail.

Although urban designers and authorities already make use of the data available to them, traditional data collecting methods tend to be more static and data is not so easy to update, whereas spatial information is exponentially more valuable when the time layer is included in the equation. Deriving the data from constantly self-updating sources, and establishing a workflow that collects, organizes, and processes this data ensures access to up to date scenarios at all times. Additionally, the accumulation of this data creates a spatio-temporal archive in time, which would illustrate the development and evolution of urban use patterns.

Through this research we have integrated different computational approaches together to create an automated workflow that draws data from online sources and visualizes them on maps. A major part of this workflow can be utilized to draw data from various other sources with minor tweaks in the query code. We present multiple layers of urban data that we believe to be relevant for aiding urban design decisions by revealing various urban use patterns and distribution of public and commercial amenities and services within a neighborhood. We visualize our data layers on maps which we propose to constitute a basis for further study where geospatial statistical analyses will be performed to achieve quantitative results.

We believe that with the further development of web technologies and integration of mobile devices into our lives, the definition of a city and what it encompasses is already being redefined fundamentally. Therefore, approaches to urban analysis and design must be re-evaluated and become as dynamic our urban environments already have.

Acknowledgements

The authors would like to thank Can Kadir Sucuoğlu and Cemal Koray Bingöl for their outstanding contributions in this study.

References

Balaban, O., & Tuncer, B. (2016). Visualizing Urban Sports Movement. In *Complexity & Simplicity - Proceedings of the 34th eCAADe Conference* (Vol. 2, pp. 89–94). Oululu, Finland: eCAADe and University of Oulu.

Batty, M. (2013). *Urban Informatics and Big Data.* London.

Batty, M., Axhausen, K., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., Portugali, Y. (2012). Smart cities of the future. *European Physical Journal: Special Topics*, 214, 481–518.

Bency, A., Rallapalli, S., Ganti, R., Srivatsa, M., & Manjunath, B. (2017). Beyond Spatial Auto-Regressive Models: Predicting Housing Prices with Satellite Imagery. In 2017 *IEEE Winter Conference on Applications of Computer Vision* (pp. 320–329). Santa Rosa, CA, USA: IEEE.

Boeing, G., & Waddell, P. (2016). New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings. *Journal of Planning Education and Research,* 1–20.

Chen, N. C. (2012). *Urban Data Mining: Social Media Data Analysis as a Complementary Tool for Urban Design*. MIT.

Clarke, A., & Steele, R. (2011). How personal fitness data can be re-used by smart cities. In *7th International Conference on Intelligent Sensors, Sensor Networks and Information Processing* (pp. 395–400). Adelaide, SA, Australia: IEEE.

Cohen, J. P., & Coughlin, C. C. (2008). Spatial hedonic models of airport noise, proximity, and housing prices. *Journal of Regional Science*, 48(5), 859–878.

Cranshaw, J., Schwartz, R., Hong, J. I., & Sadeh, N. (2012). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media* (pp. 58–65). Palo Alto, CA, USA: The AAAI Press.

Geoghegan, J., Wainger, L. a., & Bockstael, N. E. (1997). Analysis Spatial landscape indices in a hedonic framework:an ecological economics analysis using GIS. *Ecological Economics*, 23, 251–264.

Glaeser, E. L., Duke-Kominers, S., Luca, M., & Nalk, N. (2015). *Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life. HKS Faculty Research Working Paper Series.* Cambridge, MA.

Jacobs, J. (1961). *The Death and Life of Great American Cities.* New York, NY, USA: Random House.

Jenkins, A., Croitoru, A., Crooks, A. T., & Stefanidis, A. (2016). Crowdsourcing a collective sense of place. *PLoS ONE, 11*(4), 1–20.

Jiang, S., Alves, A., Rodrigues, F., Ferreira, J., & Pereira, F. C. (2015). Mining point-of-interest data from social networks for urban land use classification and disaggregation. *Computers, Environment and Urban Systems*, *53*, 36–46.

Kurgan, L. (2013). Million-Dollar Blocks. In *Close up at a distance: mapping, technology, and politics*. Zone Books.

Lee, J.-G., & Kang, M. (2015). Geospatial Big Data: Challenges and Opportunities. *Big Data Research*, 2(2), 74–81.

Li, S., Ye, X., Lee, J., Gong, J., & Qin, C. (2017). Spatiotemporal Analysis of Housing Prices in China: A Big Data Perspective. *Applied Spatial Analysis and Policy*, 10(3), 421–433.

Niederer, S., Colombo, G., Mauri, M., & Azzi, M. (2015). Street-level City Analytics: Mapping the Amsterdam Knowledge Mile. In *Hybrid City 2015: Data to the People* (pp. 215–220). URIAC.

Pettit, C., Widjaja, I., Russo, P., Sinnott, R., Stimson, R., & Tomko, M. (2012). Visualisation support for exploring urban space and place. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, I-2*, 153–158.

Schreck, T., & Keim, D. (2013). Visual Analysis of Social Media Data. *Computer, 46*(5), 68–75.

Peng, D., Biagi, L., & Ito, U. (2016). Leading countries based on number of monthly active Instagram users as of 1st quarter 2016. Retrieved October 5, 2017, from http://www.statista.com

Townsend, A. M. (2013). *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. New York: W. W. Norton & Company Inc.

Waddell, P., Berry, B. J. L., & Hoch, I. (1993). Residential property values in a multinodal urban area: New evidence on the implicit price of location. *The Journal of Real Estate Finance and Economics, 7*(2), 117–141.

Yemeksepeti'nden 2016 lezzet almanağı. (2016). Retrieved October 5, 2017, from http://www.yemeksepeti.com/