DIGITAL SOCIETY AS A DETERMINING FACTOR IN MOBILITY, URBAN DYNAMICS AND CURRENT CITIES STRUCTURE

Victor Manuel Castillo Luna
PhD Student UPM, Madrid

ABSTRACT

The development and evolution of the digital society has brought about relevant changes in the way people experience the city today. In particular, its impact on mobility patterns is evident, especially in the Post-COVID-19 society, where it is possible to access different types of daily activities online (e.g. working, shopping, training, etc.). At the same time, this digital era brings with it new challenges for urban and transport planning, in which information and communication technologies (ICT) can play a relevant role as transformers of traditional urban structures. For this reason, studying and understanding how the digital society is transforming current cities is important and necessary to optimize planning patterns/guidelines and achieve more sustainable and inclusive cities.

To address this challenge, this project has implemented multivariate statistical techniques to characterize the city of Madrid (Spain) at urban level, in order to detect the impact of ICTs on the current urban structure of the city. The methodology incorporated both factorial and cluster analysis with spatial restriction, based on a set of starting indicators referring to urban equipment, demography, socioeconomics, transportation, as well as variables of the digital society (e.g., electronic mailboxes-online commerce, free public and commercial Wi-Fi points, etc.).

The results, among other things, show an optimal classification grouping, as well as the way it was obtained. In this case, the optimal classification grouping is the one conformed by 6 clusters, in which different relationships were detected between the space built, the equipment and aspects of the digital society such as online shopping and the proximity to transport equipment public and physical commerce. These results support a growing influence of the digital society in the urban characterization, suggesting the need to translate this influence into effective planning criteria.

1. INTRODUCTION

The emergence of the digital era in society triggered an unprecedented explosion in the multiple ways in which people experience the city, giving a new direction to its development and modifying urban dynamics, mobility and spaces (Batty et al., 2012). In this sense, the UN´s Agenda 2030 considers digital tools as an essential means in the study and understanding of cities, as well as to promote socioeconomic development, protect the environment, modernize infrastructure and improve human progress (United Nations, 2015). For this reason, there is a tangible need for developing and implementing innovative solutions that use criteria related to digital society in the analysis of the urban form.

The notion of digital society proposes that network infrastructures, as well as information and communication technologies (ICTs), are capable of generating a large amount of data (Butot et al., 2020) and facilitating development processes. Therefore, they can be used to provide solutions to social, economic, urban and territorial challenges (Townsend, 2013) , which guide the progress and growth of cities. On the other hand, urban characterization, defined by its capacity to identify and classify urban typologies through various attributes related to the city, allows the elaboration of quantitative analyses with multiple variables (principal component analysis, factorial, cluster, etc.), using a large amount of information (Song & Knaap, 2007). As a whole, digital society and urban characterization, promote the idea that "increasingly complex" urban processes (Kitchin, 2016), can be made understandable by integrating both concepts. However, ICTs bring new challenges in the planning and development of cities, because by themselves, these new technologies aren´t able to solve problems in a simplistic way (Barbosa et al., 2020). This also generates several inconveniences (e.g. accumulation of power of those who generate and manage a high volume of data) (Sadin, 2013), therefore, their role is and will be fundamental as transformers of traditional urban structures (Elias B., 2020).

Although the impact of the digital society on the structure of cities has been an area of interest in academia in recent years, (Batty et al., 2012; Butot et al., 2020; Elias B., 2020; Reddy & Reinartz, 2017; Ricaurte-Quijano et al., 2017; Townsend, 2013) it has not been widely explored from perspectives of urban characterization through multivariate statistical methods. In previous researches on urban issues and mobility, characterization methodologies have been developed using multivariate analysis that have ignored issues related to the digital society. These methodologies have focused mainly on indicators related to the population structure and its relationship to the metropolitan phenomenon (Martori & Hoberg, 2008; Santos P., 1991; Teoh et al., 2020; Yeh et al., 1995). Other studies have used variables related to the urban built environment (Berrigan et al., 2010; Porta et al., 2006; Song & Knaap, 2007; Vandenbulcke et al., 2009). While some other authors have developed methodologies of multivariate characterization using indicators related to attitudes and urban activities of the population (Balram & Dragićević, 2005; Jacques & El-Geneidy, 2014; Jiang et al., 2012; Steiger et al., 2015). Finally, some have even incorporated attributes with the ability to measure social welfare (Romillo B, 2013).

Furthermore, the influence of digitalization on society has been analyzed from various points of view. The first of these, focusing on the availability and access of the population to the internet, from the logic of those who have it and those who do not (Chen, 2012; El Colegio de la Frontera Norte, 2018; *Schleife*, 2010). The digital society has also been studied through the implementation of ICT´s in the identification of urban activity patterns (location of places of shopping, leisure, employment, study, etc.) using data from social networks (Steiger et al., 2015). Moreover, many studies have been developed without criteria of contiguity, obtaining significant results in relation to the attributes used, but maintaining high levels of fragmentation in these results.

Therefore, the importance of urban characterization methodologies with the capacity to incorporate attributes related to the information society is fundamental. The objective of this study is to apply a quantitative method able to characterize through a multivariate and spatial analysis, the urban typologies affected by the digital society existing in the metropolitan area of the municipality of Madrid. For this purpose, sociodemographic and socioeconomic indicators were used, as well as variables of the built urban space and, specifically, criteria related to ICT´s and spatiality.

2. CONCEPTUAL FRAMEWORK

(Song & Knaap, 2007) emphasize the usefulness of identifying and classifying urban typologies, because this facilitates understanding by organizing ideas and characteristics into defined elements. Likewise, it allows the elaboration of quantitative analyses using a large amount of information, grouping it into a set of components. They also highlight the need to characterize these areas for the development and implementation of public policies. At the same time, (Wu & Sharma, 2012), (Song & Knaap, 2007), agree that in general terms, the classifications referring to the conceptualization of urban areas, are divided into two major categories.

2.1. Descriptive classification

The first category is known as descriptive, and is mainly based on spatial divisions that are readily available or prepared from visual interpretations of maps and images (Jones, Leishman & Watkins, 2005; Thibodeau & Goodman, 2007; Watkins, 2001). Satellite images have played a key role in descriptive characterization and have been used in various approaches. Tapiador, Avelar & Tavares-Correa (2011) applied this type of images to detect the urban morphology of an area of Lima, Peru and to classify it by relating this information to socioeconomic aspects, this allowed them to characterize large areas in social terms and map social inequality. Jun, Jinmei, Guoyu & Jizhong (2011) characterized an urban-rural region of Qingdao with the help of high-resolution satellite images under the criterion of land use properties (such as buildings, roads, pastures, farmland and water).

This type of descriptive characterization has been used in studies of the real estate market to spatially classify housing sub-markets under generic attributes, such as the census blocks into which the dwellings are grouped (Thibodeau & Goodman, 2007), postcodes (Jones et al., 2005) or the physical characteristics of the houses (Basu & Thibodeau, 1998; Watkins, 2001). However, there is a notable lack of descriptive characterizations based on attributes of the digital society (e.g. proximity of the houses to places with public internet access).

2.2. Multivariate classification

The second category is based on the interaction of data from different variables, where the growth of GIS tools and technical innovations has allowed statisticians to process a wide variety of quantitative attributes simultaneously. There are two variations in this category: without and with spatial restriction criteria.

2.2.1 Multivariate classification without spatial criteria

This first modality allows vast amounts of information and attributes to be covered without considering elements of contiguity, so the urban structure can be characterized from different perspectives, although not in a localized manner.

One perspectives is through the urban structure of the population and its relationship with the metropolitan phenomenon using socioeconomic variables and educational level, age indices, diversity and habitability indicators, and using factor analysis (FA), principal components analysis (PCA) and cluster analysis (CA). This reveals that the oldest and most educated population lives in the areas with the highest socioeconomic level (Martori & Hoberg, 2008; Santos P., 1991; Teoh, Anciaes & Jones, 2020; Yeh et al., 1995).

Likewise, urban characterizations have been studied using the built environment through building type indicators (size of homes, apartments and building, number of levels, building density, etc.), section widths, accessibility to public transport, etc. (Berrigan et al., 2010; Porta et al., 2006; Song & Knaap, 2007; Vandenbulcke et al., 2009). This group does not included any built-environment attribute related to the digital society either.

City characterizations have been developed with attributes associated with the population´s activities (work, study, leisure, etc.) and travel behavior using origin and destination indicators, travel distances, schedules, mode of transport, activities during the trip, etc. (Balram & Dragićević, 2005; Ettema & Verschuren, 2008; Jacques & El-Geneidy, 2014; Jiang et al., 2012; Steiger et al., 2015; Varghese et al., 2018). This type of study analyzes the relationship between the digital society and travel behaviour, but with the aim of identifying the influence of digital media on multitasking and rather than as an attribute for effectively contributing to urban characterization.

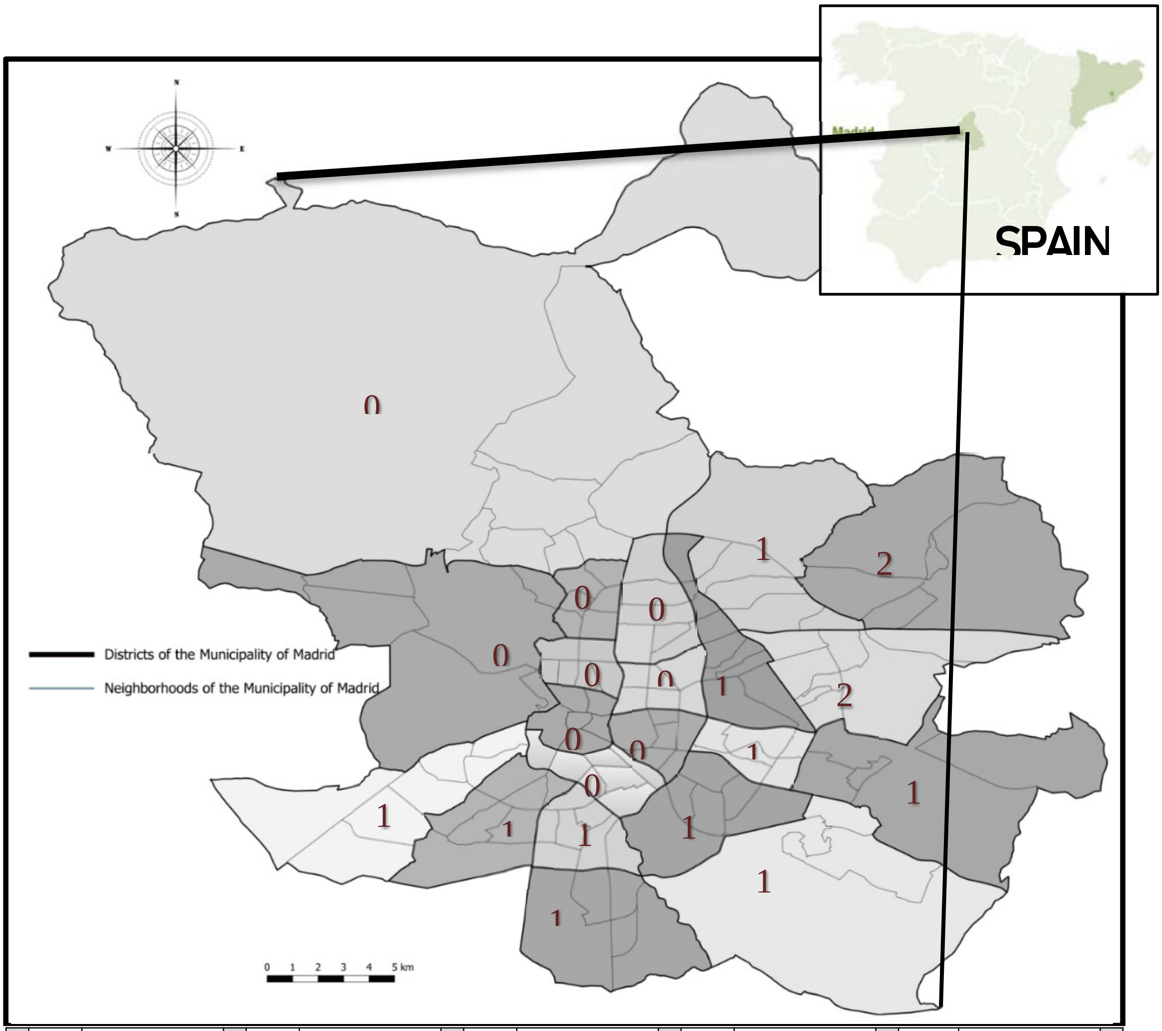
Some authors have used social indicators (percentage of active population and rates of: unemployment, divorce, infant mortality, life expectancy, etc.) in PCA and CA to group the sectors of an urban area according to territorial structures of social welfare (Romillo B, 2013), but without including parameters related to the digital society in their classifications.

Characterization with this technique has also been extremely useful in explaining the phenomenon or urban expansion, identifying sub-centers in urban areas (Cai, Huang & Song, 2017; Thomas, Riguelle & Verhetsel, 2007) and studying the effects of the compact vs. dispersed city (Thinh, Arlt, Heber, Hennersdorf & Lehmann, 2002). Or simply studying the phenomenon of sprawl by characterization the structure of the built environment and its connection with the growth of cities (Batty, Longley & Fotheringham, 1989; Piron, Dureau & Mullon, 2004).

2.2.2 Multivariate classification with spatial criteria

In the second mode, spatial restrictions can be incorporated simultaneously to the data being classified using criteria of location, distance or contiguity in the variables, or by including additional indicators that restrict the spatiality of the results.

Many studies incorporating contiguity restrictions have been undertaken in the real estate market, and include characteristics related to dimensions, costs, building heights, number of rooms etc. This allows similar areas to be identified and grouping geographically (Bates, 2006; Bourassa et al., 2007; Clapp & Wang, 2006; Tu, Sun & Yu, 2007; Wilhelmsson, 2004).



DIST.	NEIGHBORHOOD
01 CENTRO	011 Palacio 012 Embajadores 013 Cortes 014 Justicia 015 Universidad 016 Sol
02 ARGANZUELA	021 Imperial 022 Acacias 023 Chopera 024 Legazpi 025 Delicias 026 Palos de Morguer 027 Atocha
03 RETIRO	031 Pacífico 032 Adelfas 033 Estrella 034 Ibiza 035 Jerónimos 036 Niño Jesús
04 SALAMANCA	041 Recoletos 042 Goya 043 Fuente del Berro 044 Guindalera 045 Lista 046 Castellana
05 CHAMARTÍN	051 El Viso 052 Prosperidad 053 Ciudad Jardín 054 Hispanoamérica 055 Nueva España 056 Castilla
06 TETÚAN	061 Bellas Vistas 062 Cuatro Caminos 063 Castillejos 064 Almenara 065 Valdeacederas 066 Berruguete
07 CHAMBERÍ	071 Gaztambide 072 Arapiles 073 Trafalgar 074 Almagro 075 Ríos Rosas 076 Vallehermoso
08 FUENCARRAL-EL PARDO	081 El Pardo 082 Fuentelarreina 083 Peña Grande 084 El Pilar 085 La Paz 086 Valverde 087 Mirasierra 088 El Goloso
09 MONCLOA-ARAVACA	091 Casa de Campo 092 Argüelles 093 Ciudad Universitaria 094 Valdezarza 095 Valdemarín 096 El Plantío 097 Aravaca
10 LATINA	101 Los Cármenes 102 Puerta del Ángel 103 Lucero 104 Aluche 105 Campamento 106 Cuatro Vientos 107 Las Águilas
11 CARABANCHEL	111 Comillas 112 Opañel 113 San Isidro 114 Vista Alegre 115 Puerta Bonita 116 Buenavista 117 Abrantes
12 USERA	121 Orcasitas 122 Orcasur 123 San Fermín 124 Almendrales 125 Moscardó 126 Zofío 127 Pradolongo
13 PUENTE DE VALLECAS	131 Entrevías 132 San Diego 133 Palomeras Bajas 134 Palomeras Sureste 135 Portazgo 136 Numancia
14 MORATALAZ	141 Pavones 142 Horcajo 143 Marroquina 144 Media Legua 145 Fontarrón 146 Vinateros
15 CIUDAD LINEAL	151 Ventas 152 Pueblo Nuevo 153 Quintana 154 Concepción 155 San Pascual 156 San Juan Bautista 157 Colina 158 Atalaya 159 Costillares
16 HORTALEZA	161 Palomas 162 Piovera 163 Canillas 164 Pinar del Rey 165 Apóstol Santiago 166 Valdefuentes
17 VILLAVERDE	171 San Andrés 172 San Cristóbal 173 Butarque 174 Los Rosales 175 Los Ángeles
18 VILLA DE VALLECAS	181 C.H. de Vallecas 182 Santa Eugenia 183 Ensanche de Vallecas
19 VICÁLVARO	191 C. H. de Vicálvaro 192 Valdebernardo 193 Valderrivas 194 El Cañaveral
20 SAN BLAS-CANILLEJAS	201 Simancas 202 Hellín 203 Amposta 204 Arcos 205 Rosas 206 Rejas 207 Canillejas 208 El Salvador
21 BARAJAS	211 Alameda de Osuna 212 Aeropuerto 213 C. H. de Barajas 214 Timón 215 Corralejos

Figure 1. - Area of study – Municipality of Madrid

However, there are very few studies with the capacity to characterize urban areas incorporating spatial restrictions, and there continue to be considerable gaps in the methodology. As mentioned by Wu & Sharma (2012), authors that include spatial contiguity in their methodologies (Bates, 2006; Bourassa et al., 2007; Clapp & Wang, 2006; Tu et al., 2007; Wilhelmsson, 2004; Wu & Sharma, 2012), were successful in achieving some levels of restriction, although the groupings generated lack contiguous boundaries and their results didn´t coincide with the actual division of urban space. In addition, by focusing on spatial contiguity, the studies fail to highlight the importance of the attributes analyzed with respect to the complexity of the urban space (Wu, Wei & Li, 2019). This complexity is often addressed through features alien to the digital society, leading to significant gaps when incorporating spatial criteria and information society indicators.

This highlights the need for a feasible solution for the classification of urban areas, and suggests that an optimal procedure would strike a balance between the importance of emphasizing the required urban attributes (of the built space, the conditions and habits of the population and incorporating digital society variables) and including spatial restriction criteria. This balance would make it possible to achieve a homogeneous classification without mitigating the original attributes.

3. METHOD AND RESEARCH DESIGN

3.1. Case of study

The municipality of Madrid, in the metropolitan area of Madrid, Spain, was selected for this study (Fig. 1), as it is considered to fulfil many of the characteristics that exemplify the influence of the use of new ICT´s in Spain and Europe.

According to the survey *“Equipamiento y Uso de Tecnologías de la Información y Comunicación en los Hogares de la Comunidad de Madrid. TIC-H(2019a)”* 94.5% of the homes in this community had access to internet, and it is also in second place in the province of Madrid, in terms of internet users with 94.1% of people connected (Instituto Nacional de Estadística, 2019b).

In addition to this, a digital transformation has been developed in various areas of the urban structure of Madrid, with transport being one of the main ones by focusing its attention on four lines of action. The traveler´s experience, security, intelligent platforms and routes (automating a wide variety of services, and offering digitally information related to trips, waiting times, routes, etc.) and finally from the line of environmental impact, addressing urban space management strategies to reduce congestion and environmental pollution (Public Sector Observatory, 2018).

3.2. Description and data

Our analysis was performed using indicators from the municipality of Madrid, created according to the urban and digital characteristics of the study area, and classified into six groups as shown in Table 1.

VARIABLES			
GROUPS	**VARIABLE**	**No.**	**UNIT**
1.- Socio-demographic variables	Housing Density	**1**	(Inhab / dw)
	Employment Density	**2**	(Inhab/Emp)
	% of population without education	**3**	%
	% of population with "University Studies"	**4**	%
	% of population dedicated to the "Industry" sector	**5**	%
	% of population dedicated to the "Construction" sector	**6**	%
	% of population dedicated to the "Services" sector	**7**	%
	Housing rental levels by neighborhood	**8**	(Euros /m2)
	Housing price levels by neighborhood	**9**	(Euros /m2)
2.- Diversity Variables	Youth Index	**10**	Index
	Ageing Index	**11**	Index
	Index: Structure of the working population	**12**	Index
	Shannon Diversity Index	**13**	Index
3.- Design and urban structure variables	Density of public transport stops and stations	**14**	(Stops/1000inhab)
	Proximity of the population to public transport stops (bus, metro, suburban)	**15**	Distancia
	% of land use dedicated to "Residential"	**16**	%
	% of land use dedicated to "Green areas"	**17**	%
	% of land use dedicated to "Tertiary services"	**18**	%
	% of land use dedicated to roads (streets and sidewalks)	**19**	%
	Number of intersections per unit of analysis	**20**	No.
	Density of intersections per neighborhood	**21**	Intersec./1000inhab
4.- Physical commerce variables	Density of basic physical stores	**22**	Physical stores/1000inhab
	Proximity of the population to physical commercial activities of daily use	**23**	Distance
	Number of shopping centers per unit of analysis	**24**	No.
	Number of municipal markets per unit of analysis	**25**	No.
	Number of supermarkets per unit of analysis	**26**	No.
	Density of supermarkets per unit of analysis	**27**	Superm/1000 inhab
5.- Information society variables	Density of electronic mailboxes	**28**	Elec. Mail./1000 inhab
	Density of Free Public Wifi Points	**29**	(F.W.P. / 1000 inhab)
	Density of Commercial Public Wifi Points	**30**	(W.P.C./1000 inhab)
	Proximity of the population to electronic mailboxes	**31**	Distance
	Proximity of the population to free public wifi points per unit of analysis	**32**	Distance
	Proximity of the population to public wifi points of stores per unit of analysis	**33**	Distance
	Number of electronic mailboxes per analysis unit	**34**	No.
	Number of points with free public wifi per unit of analysis	**35**	No.
	Number of points with public wifi of stores per unit of analysis	**36**	No.
6.- Location variables	Distance to the centroid	**1**	Distance
	District of belonging	**2**	No.
	Geographic coordinates	**3**	Coordinates
	Cluster of belonging	**4**	No.

Table 1. - Variables

These variables were developed using the 131 neighborhoods in the 21 districts in the municipality of Madrid shown in Figure 1 as units of analysis. They were obtained with the help of different official databses, as well as various digital platforms (NOMECALLES, NavegaPorMadrid, GoogleMaps, QGis, etc.), websites of official national agencies or of the Madrid regional government (INE, website of the Madrid City Council and the Madrid Transport Consortium, the open data portal of the Municipality of Madrid, etc.), websites of stores, parcel services, supermarkets, tourism, digital services, among other sources.

A data based urban characterization model was applied through factor analysis (FA) and cluster analysis. This methodology without spatial contiguity parameters has been previously tested in several studies focused on urban structure and metropolitan phenomena (Jacques & El-Geneidy, 2014; Martori & Hoberg, 2008; Piron et al., 2004; Santos P., 1991; Song & Knaap, 2007; Thinh et al., 2002; Thomas et al., 2007; Yeh et al., 1995). Including criteria of contiguity and spatial constraints some of them, mainly focused on the analysis of the housing market (Bates, 2006; Bourassa et al., 2007; Clapp & Wang, 2006; Tu et al., 2007; Wilhelmsson, 2004).

4. RESULTS

4.1. Factorial analysis (FA)

The application of FA allowed seven factors to be extracted that can reproduce approximately 71% of the total variation of our original 36 variables, given their scores. Principal component analysis was used for factor extraction. Varimax with Kaiser Normalization was applied as the rotation method in the analysis, as this combination explains the greatest variation in the data by maximizing the loading squared for each component (Campos, Pitombo, Delhomme & Quintanilla, 2020; Khan, Ahmad & Bano, 2020; Wu & Sharma, 2012).

The results of the FA are shown in Table 2, where the variables are listed in the order of magnitude of their factor loadings sequentially for each factor. This shows the seven dimensions (factors) of the attributes originally proposed (socio-demographic, diversity, information society, etc.) resulting from the analysis.

The first factor reflects the dimension "*Information Society: Proximity and distribution of equipment*". It mainly highlights the variables of the information society and then, those related to public urban equipment's. The high correlation between ICT variables and public equipment's (public transportation stops and physical commerce) is a clear indicator that there aren´t physical barriers that interfere with the user´s choice to make purchases and develop activities in a traditional way or using digital media.

The second factor includes variables of “*Diversity and surplus value of urban areas*”, with high positive loads existing among the variables related to the cost of renting, the indices of the active population and of Shannon´s diversity.

The third factor reflects the “*Socioeconomic*” level. High positive loads among the variables: percentage of population with university studies (PPUNS) and housing costs per m^2, and high negative loads in the variables: percentage of population without studies (PPWS), percentage of population dedicated to the service sector (PPSS) and employment density.
The fourth factor is related to the variables of “*Density of urban and digital equipment*”. Maintaining high positive loads between the density of physical commerce, density of wifi points in stores (WPC), density of supermarkets and density of public transport stops (DPTS). Thus corroborating the capacity of this factor to represent aspects related to the public equipment’s in the analysis units.

FACTORIAL ANALYSIS Rotated component matrix Components	1	2	3	4	5	6	7
1.-Hous_Density	0.472	-0.656					0.231
2.-Emp_Density	0.390		-0.627	-0.221		-0.505	
3.-Youth_index	0.431	-0.362					0.623
4.-Aging_index	-0.338	0.522					-0.516
5.-Pop%(Witho_stud)			-0.902				
6.-Pop%(Univ_stud)			0.952				
7.-Pop%(Indus)							
8.-Pop%(Constr)			-0.666	-0.303			
9.-Pop%(Serv)			0.506	0.202			
10.-Rent(Eu/m2)	-0.392	0.614	0.452				
11.-Cost(Eu/m2)		0.324	0.863				
12.-Act_pop_index	0.251	0.715					
13.-Shannon_Ind	-0.294	0.787		0.221	-0.204		
14.-P.T. Density		-0.346		0.719			0.203
15.-P.T. Distance	0.730				0.210		
16.- %Resid_Use	-0.360	0.521	0.203		-0.493		
17.-%Green_Areas_Use		-0.200		-0.254	0.687		
18.-%Serv_Use							0.747
19.-%Road_use	-0.243	0.375		-0.218	-0.502	-0.236	0.280
20.-Intersections	0.484				0.304	0.542	0.223
21.-Inters_Density	0.702			0.468			
22.-Phys_Com_Dens.		0.225		0.880			
23.-Phys_Com_Dist.	0.822					0.201	
24.- No.Shop._Center						0.272	0.542
25.- No. Municip_Mark		0.603					
26.- No. Superm.		0.659				0.459	
27.-Superm_Density		0.212	0.369	0.760			
28.-Mailbox_Density				0.530		0.604	
29.-F.W.P. Density				0.639	0.645		
30.-W.P.C. Density		0.239		0.819			
31.-Mailbox_distance	0.930						
32.-F.W.P. Distance	0.912						
33.-W.P.C. Distance	0.859		-0.243				
34.-No.Mailbox		0.305				0.821	
35.-No.F.W.P.		0.478			0.739		
36.-No.W.P.C.		0.641	0.335	0.316		0.225	0.207
Extraction method: principal component analysis							
a. The rotation has become 9 interactions							

Table 2. - Matrix of components/factors

The fifth factor incorporates variables of “*Digital Society as a function of Land Use*”, keeping mainly high factorial loads between the number and density of free wifi poins (FWP) and the percentage of land use dedicated to green areas (PLUGA). Highligthing the capacity of this factor to cover information regarding the sites where there is greater possibility of finding FWP according to the type of land use present in each unit of analysis.

The sixth factor reflects interesting parameters among the variables related to “*Online Shopping vs employment density*”. With high positive factorial loads among the variables: number and density of electronic mailboxes, and relatively high negative factorial load with the variable employment density.

Finally, the seventh factor includes variables related to the “*Structure of the population*”. Having high positive factorial loads between the youd index and the land use dedicated to services, with a high negative factorial load focused on the aging index. This allows us to differentiate the areas where the young population predominates, and in which of them, on the contrary, there is an aging population structure.

4.2. Cluster analysis (CA)

One of the main objectives of the research is to understand the dimension, variation and influence of the characteristics described in the previous factors incorporating spatial restrictions. It was therefore necessary to identify the optimal number of groups in which to incorporate our units of analysis. However, due to the lack of previous studies specific to the metropolitan area under the characteristics considered, a hierarchical cluster analysis was used to determine the appropriate k-value (number of groups).

With the factors extracted in FA, the variables of spatial constraints and the number of cluster (k-value) defined, six k-means cluster analysis (CA K-means) with 125 units of analysis were performed. To form clusters with the influence of spatial contiguity from smallest to largest measure, a sequence of incorporation of spatial constraints variables was assigned for each cluster analysis.

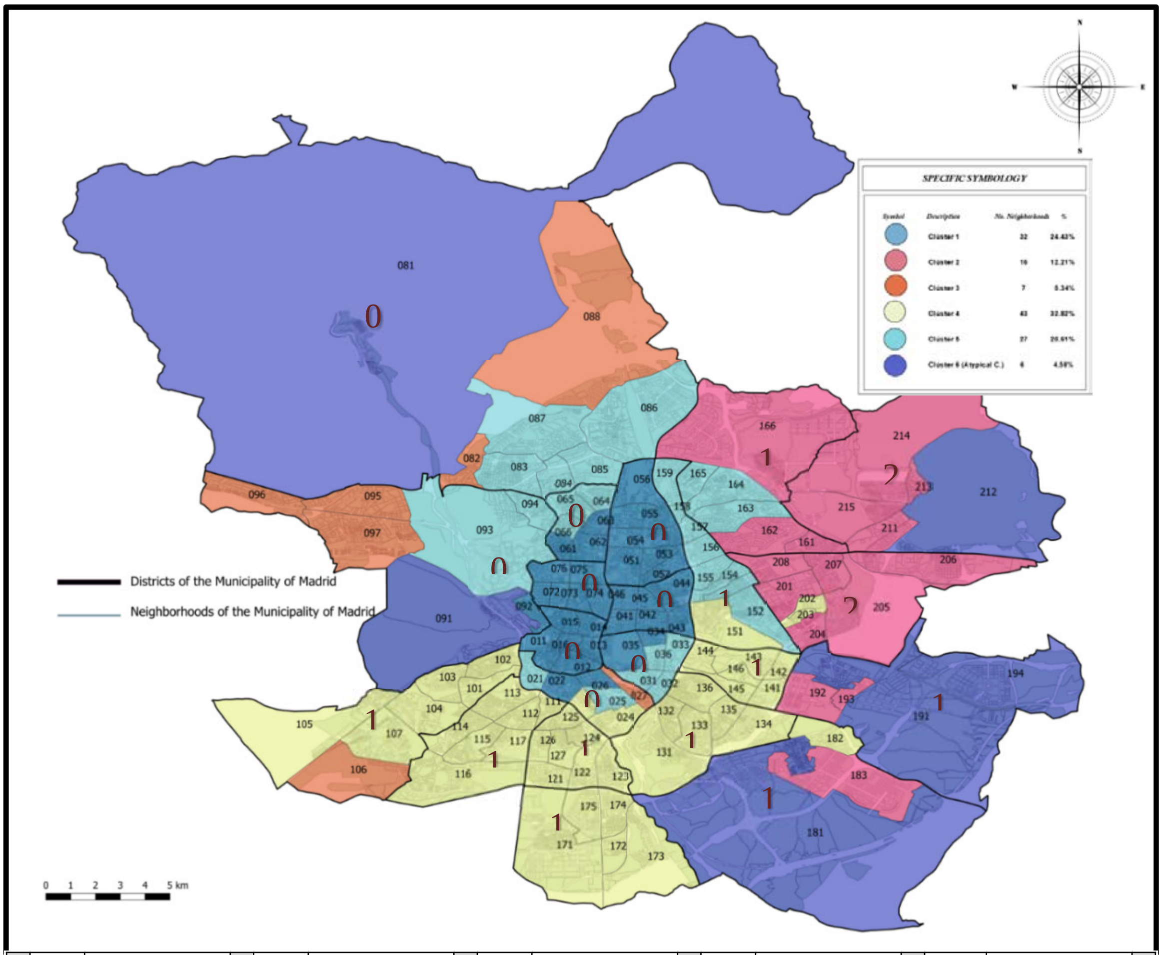
Table 3 shows the number of cases grouped in the clusters formed. It was observed that the number of cases is more homogeneously distributed in the clusters with a greater number of spatial constraints variables.

The results were reviewed to select the optimal classification method according to the objective of the study. The result was that Group V, comprising seven FA factors and four spatial constraint variables, is the cluster that best defines the structure of the metropolitan area of the municipality of Madrid according to the characteristics considered (Fig. 2).

Clúster #	7 Factors Group I	7 Factors + 1 Var. Restriction Group II	7 Factors + 2 Var. Restriction Group III	7 Factors + 3 Var. Restriction Group IV	7 Factors + 4 Var. Restriction Group V	Constraint Spatial "K-means clustering" Group VI
1 (Blue)	50	58	53	47	32	12
2 (Pink)	17	17	15	24	16	34
3 (Orange)	3	2	4	2	7	20
4 (Green)	53	46	51	45	43	38
5 (Light blue)	2	2	2	7	27	21
Subtotal	**125**	**125**	**125**	**125**	**125**	**125**
6 (Atypical cases)	6	6	6	6	6	6
Total	**131**	**131**	**131**	**131**	**131**	**131**

Table 3. - K-means cluster analysis (K = 5)

The map represents the 131 units of analysis (Fig. 2). Except for the atypical cases that make up cluster six, the formation of homogenous sub-groups can be seen with units that maintain a contiguous boundary with another unit belonging to the same group, or at least have a location close to the subgroup. Achieving with this, an adequate level of spatial integrity and a balance between the location parameters used, the characteristics of urban structure and the variables of digital society considered (Table 4).



DIST.	NEIGHBORHOOD
01 CENTRO	011 Palacio 012 Embajadores 013 Cortes 014 Justicia 015 Universidad 016 Sol
02 ARGANZUELA	021 Imperial 022 Acacias 023 Chopera 024 Legazpi 025 Delicias 026 Palos de Morguer 027 Atocha
03 RETIRO	031 Pacífico 032 Adelfas 033 Estrella 034 Ibiza 035 Jerónimos 036 Niño Jesús
04 SALAMANCA	041 Recoletos 042 Goya 043 Fuente del Berro 044 Guindalera 045 Lista 046 Castellana
05 CHAMARTÍN	051 El Viso 052 Prosperidad 053 Ciudad Jardín 054 Hispanoamérica 055 Nueva España 056 Castilla
06 TETÚAN	061 Bellas Vistas 062 Cuatro Caminos 063 Castillejos 064 Almenara 065 Valdeacederas 066 Berruguete
07 CHAMBERÍ	071 Gaztambide 072 Arapiles 073 Trafalgar 074 Almagro 075 Ríos Rosas 076 Vallehermoso
08 FUENCARRAL-EL PARDO	081 El Pardo 082 Fuentelarreina 083 Peña Grande 084 El Pilar 085 La Paz 086 Valverde 087 Mirasierra 088 El Goloso
09 MONCLOA-ARAVACA	091 Casa de Campo 092 Argüelles 093 Ciudad Universitaria 094 Valdezarza 095 Valdemarín 096 El Plantío 097 Aravaca
10 LATINA	101 Los Cármenes 102 Puerta del Ángel 103 Lucero 104 Aluche 105 Campamento 106 Cuatro Vientos 107 Las Águilas
11 CARABANCHEL	111 Comillas 112 Opañel 113 San Isidro 114 Vista Alegre 115 Puerta Bonita 116 Buenavista 117 Abrantes
12 USERA	121 Orcasitas 122 Orcasur 123 San Fermín 124 Almendrales 125 Moscardó 126 Zofío 127 Pradolongo
13 PUENTE DE VALLECAS	131 Entrevías 132 San Diego 133 Palomeras Bajas 134 Palomeras Sureste 135 Portazgo 136 Numancia
14 MORATALAZ	141 Pavones 142 Horcajo 143 Marroquina 144 Media Legua 145 Fontarrón 146 Vinateros
15 CIUDAD LINEAL	151 Ventas 152 Pueblo Nuevo 153 Quintana 154 Concepción 155 San Pascual 156 San Juan Bautista 157 Colina 158 Atalaya 159 Costillares
16 HORTALEZA	161 Palomas 162 Piovera 163 Canillas 164 Pinar del Rey 165 Apóstol Santiago 166 Valdefuentes
17 VILLAVERDE	171 San Andrés 172 San Cristóbal 173 Butarque 174 Los Rosales 175 Los Ángeles
18 VILLA DE VALLECAS	181 C.H. de Vallecas 182 Santa Eugenia 183 Ensanche de Vallecas
19 VICÁLVARO	191 C. H. de Vicálvaro 192 Valdebernardo 193 Valderrivas 194 El Cañaveral
20 SAN BLAS-CANILLEJAS	201 Simancas 202 Hellín 203 Amposta 204 Arcos 205 Rosas 206 Rejas 207 Canillejas 208 El Salvador
21 BARAJAS	211 Alameda de Osuna 212 Aeropuerto 213 C. H. de Barajas 214 Timón 215 Corralejos

Fig. 2. - Group V – 4 Spatial constraints variables (Best grouping)

CLUSTER´S	
CLUSTER	**Main features**
1	Central zone with ederly population
	High ICT services (electronic mailboxes, Wifi Points Commerce)
	High urban services (public transport stops, basic retail outlets, municipal markets)
	High rent and housing cost and high socioeconomic level
	High percentage of population with university studies
	Shorter travel distances to access digital society services
	Lower number of free wifi points.
2	Northeast zone with young population
	High number of electronic mailboxes
	Low density of public services and low active population and Shannon diversity indices
3	Mixed zone with High percentage of land use dedicated to green areas
	High number of free wifi points
	Longer travel distances to acces digital society services
	Structure population is mainly young
	Electronic mailboxes is relatively low
4	Southwest zone with low socioeconomic level and ederly population
	Minimum density of urban and digital services
	Low quantity of electronic mailboxes
	Low rental and housing costs with high percentage of population without studies (PPWS)
5	Peripheral zone to the center with few free wifi points and green areas
	Low density of urban and digital services
	Relatively low travel distances to access digital society services
	Above-average diversity and surplus value according to the high values of the Shannon and active population indices.
	High rental and housing costs.
6	These units are that they are located on the periphery of the study area,
	Urbanized land of less than 50%, few public and digital amenities

Table 4. - Clusters formed

5. CONCLUSIONS

The multivariate and spatial analysis developed allowed the identification of the typologies existing in the metropolitan area of Madrid based on sociodemographic, socioeconomic, urban structure and diversity criteria, and specifically relating to the digital society. A classification methodology was successfully developed with two levels of spatial restrictions: the first was incorporated into the FA, where it reduced the dimension of 36 attributes in seven factors that explained over 70% of the total data; and the second involved four more contiguity variables that were gradually incorporated in the cluster analysis.

The results suggest that 125 of the 131 units of analysis considered can be assigned to one of the five typologies, while the six remaining units were classified in a specific group of atypical cases due to their extreme characteristics. The largest group comprises 43 units of analysis (Cluster 4) and includes neighborhoods with a low socioeconomic level, minimum density of urban and digital services, and a relatively low number of electronic mailboxes, meaning that they have limited access to both physical and online shopping.

The second largest group consisted of 32 units of analysis (Cluster 1) located in the central zone. They have an elderly population and a high level of digital and urban services, with the shortest distances to these facilities. The third largest cluster comprises 27 units of analysis (Cluster 5), and contains the neighborhoods with a low density of urban and digital services located on the periphery of the central zone, thus maintaining a significant diversity and surplus value with high rental and housing costs and a representative number of Wi-Fi point commerce (WPC).

The fourth (Cluster 2) and fifth (Cluster 3) clusters in order of size included 16 and 7 units of analysis respectively. Cluster 2 contains neighborhoods with a young population structure, a high number of electronic mailboxes and low densities of public amenities, while Cluster 3 is a mixed zone with extensive green areas and free Wi-Fi points (FWP), and a low proportion of housing areas.

Although the central area is closer to digital society services, this is not the area with the most digital public amenities; these are cluster 3 and 2, which have a greater number of FWP in relation to the population. Likewise, cluster 2 and 5, despite not having an high number of electronic mailboxes like other groups, maintain a considerable proportion of these facilities, pointing to the importance of online shopping in these areas, even though Group 5 has a mainly elderly population and Group 2 has a high percentage of population without studies (PPWS).

The results show the potential usefulness of obtaining a classification of urban areas using attributes of urban complexity, with spatial limits and restrictions that do not significantly affect the quality of the data, and of measuring this complexity through the structure and services of the areas analyzed. This characterization also offers a detailed view of how digital society attributes strongly influence cities, and how these indicators are capable of modifying urban structures, the inhabitants' activities and behavior, and the general development of the urban form.

Public and private organizations can use classifications of this kind to carry out/adapt urban development plans and identify specific urban areas. This methodology can also be updated for various periods with current information, or with data from previous years to provide knowledge of the urban structure in specific years. As mentioned by Song & Knaap (2007), these classifications can be used in regression analysis to select sampling strategies, for the specific design of urban areas, and to measure progress in the implementation of urban development plans that incorporate information society criteria.

REFERENCES

BALRAM, S., & DRAGIĆEVIĆ, S. (2005). Attitudes toward urban green spaces: Integrating questionnaire survey and collaborative GIS techniques to improve attitude measurements. Landscape and Urban Planning, 71(2-4), 147-162. https://doi.org/10.1016/S0169-2046(04)00052-0

BARBOSA, A. C., MORAES, T. M., TESIMA, D. T., PONTES, R. C., & LIMA, A. Z., B.Z. (2020). Smart Planning: tools, concepts and approaches for a sustainable digital transformation. https://doi.org/10.1007/978-3-030-12255-3_14.

BASU, S., & THIBODEAU, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. The Journal of Real Estate Finance and Economics, 17, 61-85.

BATES, L. K. (2006). Does neighborhood really matter? Comparing historically defined neighborhood boundaries with housing submarkets [Article]. Journal of Planning Education and Research, 26(1), 5-17. https://doi.org/10.1177/0739456X05283254

BATTY, M., AXHAUSEN, K. W., GIANNOTTI, F., POZDNOUKHOV, A., BAZZANI, A., WACHOWICZ, M., OUZOUNIS, G., & PORTUGALI, Y. (2012). Smart cities of the future. The European Physical Journal Special Topics, 214 (1), 481-518. https://doi.org/ 10.1140/epjst/e2012-01703-3

BATTY, M., LONGLEY, P., & FOTHERINGHAM, A. (1989). Urban Growth and Form: Scaling, Fractal Geometry, and Diffusion-Limited Aggregation. Environment and Planning A, 21(11), 1447-1472. https://doi.org/ 10.1068/a211447

BERRIGAN, D., PICKLE, L., & DILL, J. (2010). Associations between street connectivity and active transportation. International Journal of Health Geographics, 9(20), 20. https://doi.org/ 10.1186/1476-072X-9-20

BOURASSA, S., CANTONI, E., & HOESLI, M. (2007). Spatial Dependence, Housing Submarkets, and House Price Prediction. The Journal of Real Estate Finance and Economics, 35(2), 143-160. https://doi.org/10.1007/s11146-007-9036-8

BUTOT, V., BAYERL, P. S., JACOBS, G., & DE HAAN, F. (2020). Citizen repertoires of smart urban safety: Perspectives from Rotterdam, the Netherlands. Technological Forecasting and Social Change, 158(120164), 11. https://doi.org/https://doi.org/10.1016/j.techfore.2020.120164

CAI, J., HUANG, B., & SONG, Y. (2017). Using multi-source geospatial big data to identify the structure of polycentric cities. Remote Sensing of Environment, 202, 210-221. https://doi.org/https://doi.org/10.1016/j.rse.2017.06.039

CAMPOS, C. I., PITOMBO, C. S., DELHOMME, P., & QUINTANILHA, J. A. (2020). Comparative analysis of data reduction techniques for questionnaire validation using self-reported driver behaviors. Journal of Safety Research, 73, 133-142. https://doi.org/https://doi.org/10.1016/j.jsr.2020.02.004

CHEN, W. (2012). The Implications of Social Capital for the Digital Divides in America. The Information Society 29(1), 13-25. https://doi.org/ 10.1080/01972243.2012.739265

CLAPP, J. M., & WANG, Y. (2006). Defining neighborhood boundaries: Are census tracts obsolete? Journal of Urban Economics, 59(2), 259-284. https://doi.org/https://doi.org/10.1016/j.jue.2005.10.003

COBO, E., CORTÉS, J., & GONZÁLEZ, J. A. (2014). Prueba de significación y contraste de hipótesis. Bioestadística para no estadísticos - Universitat Politécnica de Catalunya (UPC).

CÓDIGO DEL BONUS, C. (2019). Los españoles y los códigos promocionales

DPDgroup. (2018). E-Shopper Barometer 2018. Informe España.

EL COLEGIO DE LA FRONTERA NORTE, M. (2018). Brecha digital, uso frecuente y aprovechamiento de Internet en México. Convergencia, Revista de Ciencias Sociales, Num. 79, 1-27. https://doi.org/https://doi.org/ 10.29101/crcs.v0i79.10332

ELIAS B., S. (2020). Compact Urbanism and the Synergic Potential of its Integration with Data-Driven Smart urbanism: An Extensive Interdisciplinary Literature Review. Land Use Policy, 97(104703). https://doi.org/https://doi.org/10.1016/j.landusepol.2020.104703

ETTEMA, D., & VERSCHUREN, L. (2008). Multitasking and Value of Travel Time Savings. Transportation Research Record Journal of the Transportation Research Board 2010, 19-25. https://doi.org/ 10.3141/2010-03

IABeCommerce - ELOGIA, I. E. (2020). Estudio eCommerce 2020. https://iabspain.es/estudio/estudio-anual-de-ecommerce-2020/

INSTITUTO NACIONAL DE ESTADÍSTICA, I. (2019a). "Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares, TIC-H 2019", España". https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608

INSTITUTO NACIONAL DE ESTADÍSTICA, I. (2019b). España. https://www.ine.es/

JACQUES, C., & EL-GENEIDY, A. (2014). Does travel behavior matter in defining urban form? A quantitative analysis characterizing distinct areas within a region. Journal of Transport and Land Use, 7(1), 1-14. https://doi.org/10.5198/jtlu.v0i0.377

JIANG, S., FERREIRA, J., & GONZALEZ, M. C. (2012). Clustering daily patterns of human activities in the city. Data Mining and Knowledge Discovery 25 (3)(3), 478-510. https://doi.org/10.1007/s10618-012-0264-z

JONES, C., LEISHMAN, C., & WATKINS, C. (2005). Housing market processes, urban housing submarkets and planning policy. The Town planning review, 76 (2), 215-233. https://doi.org/10.3828/tpr.76.2.6

JUN, X., JINMEI, L., GUOYU, W., & JIZHONG, L. (2011). The classification of land cover derived from high resolution remote sensing imagery. 4th International Conference on Intelligent Computation Technology and Automation (ICICTA), 2. https://doi.org/10.1109/ICICTA.2011.472

KHAN, S. A., AHMAD, S., & BANO, R. (2020). Factorial structure of ICT components influencing consumer purchase decision in food store. International Journal of Agricultural and Statistical Sciences, 16 (1)(1), 431-438.

KITCHIN, R. (2016). The ethics of smart cities and urban science. Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences, 374(2083), 1-15. https://doi.org/10.1098/rsta.2016.0115

LOPEZ A., M., & GUTIERREZ P., L. (2019). COM DUR A TERME I INTERPRETAR UNA ANÀLISI FACTORIAL EXPLORATÒRIA UTILITZANT SPSS. REIRE: revista d'innovació i recerca en educació, 12, Nº. 2, 2019, 1-14. https://doi.org/https://doi.org/10.1344/reire2019.12.227057

MACQUEEN, J. (1967). Some methods for classification and analysis of multivariate observations University of California, Los Angeles,

MARTORI, J. C., & HOBERG, K. (2008). NUEVAS TÉCNICAS DE ESTADÍSTICA ESPACIAL PARA LA DETECCIÓN DE CLUSTERS RESIDENCIALES DE POBLACIÓN INMIGRANTE. Scripta Nova: Revista electrónica de Geografía y Ciencias Sociales, Vol. XII, núm. 263.

PIRON, M., DUREAU, F., & MULLON, C. (2004). Utilisation de typologies multi-dates pour l'analyse des transformations socio-spatiales de Bogota (Colombie). Revue européenne de géographie,, N°256, 5 février 2004. https://doi.org/https://doi.org/10.4000/cybergeo.3742

PORTA, S., CRUCITTI, P., & LATORA, V. (2006). The network analysis of urban streets: A dual approach. Physica A : Statistical Mechanics and its Applications, Volume 369, Issue 2(2), 853-866. https://doi.org/10.1016/j.physa.2005.12.063

PRINET, V., MA, S., & MONGA, O. (1999). Classification and representation of networks from satellite images. Image Analysis and Processing, 1999. Proceedings. International Conference on, 1-7. https://doi.org/10.1109/ICIAP.1999.797696

REDDY, S. K., & REINARTZ, W. (2017). Digital Transformation and Value Creation: Sea Change Ahead. Marketing Intelligence Review, 9 (1)(1), 11-17. https://doi.org/https://doi.org/10.1515/gfkmir-2017-0002

RICAURTE-QUIJANO, C., ARELLANO, A., & NARANJO, K. (2017). (Des)conexión durante el viaje turístico: uso de Smartphone por parte de viajeros independientes en la ciudad de Guayaquil. Teoría y Praxis. 23. , pp. 93-119. . https://doi.org/DOI: 10.22403/UQROOMX/TYP23/04

ROMILLO B, J. A. (2013). Métodos cuantitativos y geografía: Aplicación del análisis multivariante al estudio de los desequilibrios en materia de bienestar en España Universidad de Cantabria]. http://hdl.handle.net/10902/3862

SADIN, É. (2013). L'humanité augmentée. L'administration Numérique du Monde (Vol. ISBN: 978-2915830750.).

SANTOS P., J. M. (1991). La técnica factorial y su aplicación en el campo geográfico. Espacio, tiempo y forma. Serie VI, Geografía, Tomo 4, 1991, 13-78.

SCHLEIFE, K. (2010). What really matters: Regional versus individual determinants of the digital divide in Germany. Research Policy, Volume 39, Issue 1(1), 173-185. https://doi.org/https://doi.org/10.1016/j.respol.2009.11.003

SONG, Y., & KNAAP, G. (2007). Quantitative classification of neighbourhoods : The neighbourhoods of new single-family homes in the Portland metropolitan area. Journal of Urban Design, Vol. 12, Issue 1(1), 1-24. https://doi.org/10.1080/13574800601072640

SPEARMAN, C. H. (1904). General Intelligence," Objectively Determined and Measured (Vol. 15, No. 2). The American Journal of Psychology. https://doi.org/10.2307/1412107

STEIGER, E., RESCH, B., & ZIPF, A. (2015). Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks. International Journal of Geographical Information Science, 30 (9), 1-23. https://doi.org/10.1080/13658816.2015.1099658

STOCK, W. G. (2011). Informational cities: Analysis and construction of cities in the knowledge society. Journal of the American Society for Information Science and Technology, Volume 62, (5)(5). https://doi.org/https://doi.org/10.1002/asi.21506

TAPIADOR, F. J., AVELAR, S., TAVARES-CORREA, C., & ZAH, R. (2011). Deriving fine-scale socioeconomic information of urban areas using very high-resolution satellite imagery. International Journal of Remote Sensing, 32 (21)(21). https://doi.org/https://doi.org/10.1080/01431161.2010.512928

TEOH, R., ANCIAES, P., & JONES, P. (2020). Urban mobility transitions through GDP growth: Policy choices facing cities in developing countries. Journal of Transport Geography, 88. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102832

THE ECONOMIST - INTELLIGENCE UNIT, I. (2019). Safe Cities Index 2019. Urban security and resilience in an interconnected world. https://safecities.economist.com/safe-cities-index-2019/

THIBODEAU, T. G., & GOODMAN, A. (2007). The Spatial Proximity of Metropolitan Area Housing Submarkets. Real Estate Economics, 35(2), 209-232. https://doi.org/10.1111/j.1540-6229.2007.00188.x

THINH, N. X., ARLT, G., HEBER, B., HENNERSDORF, J., & LEHMANN, I. (2002). Evaluation of urban land-use structures with a view to sustainable development. Environmental Impact Assessment Review, 22 (5)(5), 475-492. https://doi.org/https://doi.org/10.1016/S0195-9255(02)00023-9

THOMAS, I., RIGUELLE, F., & VERHETSEL, A. (2007). Measuring Urban Polycentrism: A European Case Study and its Implications. Journal of Economic Geography, 7 (2), 193-215. https://doi.org/10.1093/jeg/lbl025

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

TREE CITIES OF THE WORLD. (2019). https://treecitiesoftheworld.org/

TU, Y., SUN, H., & YU, S. (2007). Spatial Autocorrelations and Urban Housing Market Segmentation. The Journal of Real Estate Finance and Economics, 34 (3)(3), 385-406. https://doi.org/ 10.1007/s11146-007-9015-0

UNITED NATIONS, U. (2015). Habitat III issue papers, 21—Smart cities (V2.0). Habitat III issue papers, 21—Smart cities (V2.0).

VANDENBULCKE, G., STEENBERGHEN, T., THOMAS, & I. (2009). Mapping accessibility in Belgium: a tool for land-use and transport planning? Journal of transport Geography, 17 (1)(1), 39-53. https://doi.org/10.1016/j.jtrangeo.2008.04.008

VARGHESE, V., & JANA, A. (2018). Impact of ICT on multitasking during travel and the value of travel time savings: Empirical evidences from Mumbai, India. Travel Behavior and society, 12, 11-22. https://doi.org/10.1016/j.tbs.2018.03.003

WATKINS, C. A. (2001). The Definition and Identification of Housing Submarkets. Environment and Planning A, 33, 2235-2253. https://doi.org/https://doi.org/10.1068/a34162

WILHELMSSON, M. A Method to Derive Housing Sub-Markets and Reduce Spatial Dependency. Property Management, 22 (4), 276-288. https://doi.org/ 10.1108/02637470410558143

WILHELMSSON, M. (2004). A Method to Derive Housing Sub-Markets and Reduce Spatial Dependency. Property Management, 22 (4), 276-288. https://doi.org/ 10.1108/02637470410558143

WU, C., & SHARMA, R. (2012). Housing submarket classification: The role of spatial contiguity. Applied Geography, 32 (2), 746-756. https://doi.org/https://doi.org/10.1016/j.apgeog.2011.08.011

WU, Y., WEI, Y., & LI, H. (2019). Analyzing Spatial Heterogeneity of Housing Prices Using Large Datasets. Applied Spatial Analysis and Policy 13, 223-256. https://doi.org/ 10.1007/s12061-019-09301-x

YEH, A.-O., XU, X., & HU, H. (1995). THE SOCIAL SPACE OF GUANGZHOU CITY, CHINA. Urban Geography, 16 (7)(7), 595-621. https://doi.org/https://doi.org/10.2747/0272-3638.16.7.595

YU, D., & WEI, Y. D. (2008). Spatial Data Analysis of Regional Development in Greater Beijing, China, in a GIS Environment. Papers in Regional Science, 87 (1), 97-117. https://doi.org/10.1111/j.1435-5957.2007.00148.x

ZHANG, Q., & SU, S. (2016). Determinants of urban expansion and their relative importance: A comparative analysis of 30 major metropolitans in China. Habitat International, 58, 89-107. https://doi.org/https://doi.org/10.1016/j.habitatint.2016.10.003