CLUSTERING OF HOUSEHOLDS BASED ON THEIR USE OF INFORMATION TECHNOLOGIES: A REVIEW OF TURKIYE

Research Article

Ömer Faruk RENÇBER[1]

Eda ZORKİRİŞÇİ[2]

Mehmet AYTEKİN[3]

Abstract

Access to technology is becoming increasingly available with the advancements in information and communication technologies. However, it is believed that most users now perform daily tasks on their mobile phones. In the literature, there has been a focus on examining the information and communication technology (ICT) literacy or usage status of students, teachers, and employees. However, there has been less exploration regarding the use of household ICT devices, which is a topic of interest. Today, households' ICT adaptation and usage levels are essential for creating smart buildings and cities. In this respect, the study provides an idea for people and institutions planning smart cities in Turkey.

The study categorizes household ICT devices into clusters based on their usage and examines the differences in usage patterns. The data for the year 2021 was obtained from the household information and communication technologies survey conducted by TURKSTAT in Turkey. The participants were categorized into two groups: advanced and basic users. Regarding regional differences, it is concluded that the rate of advanced users is higher in cities in western Turkey than in other regions. Based on the study, it was found that people mostly use mobile devices for daily usage. The groups differ regarding how long they use the internet, their e-commerce habits, and their behaviours when encountering website purchasing problems.

Keywords: Information and Communication, Clustering, Household, Digital Divide, Local Studies

[1] Correspond Author, Assoc. Prof., Gaziantep University Faculty of Economics and Business, dr.ofrencber@gmail.com ORCID: [0000-0001-8020-2750]

[2] PhD Student, Gaziantep University Faculty of Economics and Business, eddalbudak@gmail.com ORCID: [0000-0003-2858-5648]

[3] Assoc. Prof., Assoc. Prof., Gaziantep University Faculty of Economics and Business, aytekin@gantep.edu.tr ORCID: [0000-0001-5464-0677]

Makale Gönderim Tarihi: 7 Eylül 2023

Makalu Kabul Tarihi: 11 Ekim 2023

INTRODUCTION

Integrating information and communication technology (ICT) into our daily lives has changed how households function and interact. As technology is increasingly integrated into different aspects of life, understanding the patterns of technology adoption and use among households has become a topic of considerable interest. Following and understanding the relevant policies is essential to stay up-to-date with the latest social, cultural, and economic developments. This can help individuals make informed decisions about their actions and contribute to the overall progress of society (Susan & Brown, 2016).

The importance of digital technologies in our daily lives cannot be denied. With the increasing use of these technologies, individuals and households are experiencing an improvement in their living standards (Berg, 2022). The pandemic highlighted information technologies' significance, as physical access was limited (Perera & Abeysekera, 2022). The increased use of these technologies has led to a surge in research and studies on this subject, which is essential for the betterment of society.

The utilization of information technologies is of significant importance when it comes to making informed decisions for many individuals. Research studies often concentrate on using these technologies by students, employees, or educators (Russell et al. (2000); Hollman et al. (2019); Marler et al. (2006)). However, it is equally crucial to consider the extent of technology usage by all household members, whether desktop or mobile users. There are many studies in the literature that examine household ICT use. Some of these include technology adoption (Rogers, 1962; Rogers, 2003), the digital divide (Warschauer, 2003), social interaction and communication (Baym, 2010), education and learning (Hew & Bush, 2007), advertising and media (Vorderer et al., 2012) and changes and concerns (Kitchin, 2016).

Over time, the use of information and communication technologies has become essential. This is because they are the primary means of accessing information, which society needs. Those who want to access information quickly rely on these technologies. Additionally, society's use of this technology ensures access to reliable and accurate information (Buhalis, 1998).

By clustering households based on their use of ICT, we can gain insight into their behaviour and needs. This information can create tailored services and policies for each household. For instance, if a cluster is identified where older individuals use ICT less, education programs can be developed specifically for those households.

The data set was obtained from the household information technologies usage (HITU) study conducted by the Turkish Statistical Institute (TURKSTAT). TURKSTAT has regularly conducted the HITU since 2004 (except for 2006) to monitor usage patterns and ensure equitable access across regions. The collected data allows researchers to analyze and compare technological usage levels among

individuals within various areas, identifying disparities and offering recommendations to policymakers and business leaders.

In this study, data collected by TURKSTAT in 2021 was analyzed through cluster analysis to assign it into groups. The objective is to compare the regional usage levels of technological devices among individuals, analyze the technical usage of individuals at both regional and annual levels, and investigate differences in restricted specialized use in Turkey. Based on these aims, the paper identifies key findings and provides practical suggestions for lawmakers and company managers.

1. Method of Study

K-Means cluster analysis method was used in the study. In this section, the theoretical information of the method is given.

1.1. K-Means Method

K-Means is one of the most widely used and fundamental clustering algorithms, known for its simplicity and effectiveness. Cluster analysis is a multivariate statistical method used to divide the data into homogeneous subgroups and classify them according to their similarities. The K-Means clustering method, developed by Mac Queen in 1967, is widely recognized for its non-hierarchical structure. The basic logic of the algorithm is to divide the data set consisting of n data into k sets specified as input values. The k-means method has been widely utilized across various fields and industries, including image segmentation, market segmentation, and medical segmentation (Ng et al. (2006); Kuo et al. (2002); Dhanachandra et al. (2015)).

K-Means is an unsupervised learning algorithm designed to partition a dataset into K distinct, non-overlapping clusters. The algorithm's main objective is to minimize the sum of squared distances between data points and their respective cluster centroids. It operates under the assumption that each data point belongs to the cluster with the nearest mean (centroid).

The K-Means clustering method steps are as follows (Teknomo, 2006):

1. Initialization: Determine the k value, which is the number of clusters to be divided.

2. Assignment Step: Assign each data point to the nearest cluster centroid based on a distance metric.

3. Update: Calculate new centroids for each cluster,

4. Iteration: Step 2 and Step 3 are repeated until each data is assigned to a cluster.

The advantages of this study are simplicity, efficiency, and scalability. Moreover, limitations are initial centroid sensitivity, determining the k number, and assumption of the spherical cluster (Drias et al., 2016).

2. Variables Used in the Study

As technology becomes increasingly prevalent in households, the internet access tools are becoming more diverse. Considering these devices' impact on an individual's ability to connect to the internet is essential. In order to gain a better understanding of internet usage, a study was conducted utilizing data from TURKSTAT's HITU survey for 2021.

The study focused on the devices individuals used in the last three months to access the internet, with variables including desktop computers, laptops, tablets, cell phones or smartphones, and other technological devices such as smart TVs, smartwatches, and game consoles. By examining these variables, we can gain insight into how individuals access the internet and how this technology impacts their connectivity.

3. Implementation

The data were analyzed with the K-Means clustering method in this part of the study. This study aims to examine the differences between advanced users and basic users. The reason for choosing the number of groups is 2; This is a more balanced clustering of observations.

Table 1. Frequency of the Clusters

Clusters	Group 1	15276
	Group 2	9052
Total Participant		24328

Data obtained from 24328 individuals were used in the study. Table 1 shows that 15276 of these data are in the basic user cluster, while 9052 are in the advanced user group.

Table 2. Center Values of Groups

	Clusters	
	Group 1	**Group 2**
Desktop Computer Use of Individuals	2	2
Individuals' Use of Laptop Computers	2	1
Tablet Computer Use of Individuals	2	2
Cell Phone or Smart Phone Usage of Individuals	1	1
Other Technological Device Uses by Individuals	2	1
1: Using 2: Not Using		

Table 2 shows that Group 1 uses the internet from their mobile phones, while Group 2 uses laptop computers, mobile phones and other information devices. The

main difference between the groups is that Group 2 uses laptops and other devices (smart TV, smart speaker, game console, e-book reader, smartwatch). In contrast, the others only access the Internet through mobile devices. For this reason, group 1 is called basic users, and group 2 is called advanced users.

Table 3. Comparison of Groups by Gender

Clusters		**Frequency**	**Percent**
Basic User	Man	7892	51,7
	Woman	7384	48,3
	Total	15276	100,0
Advanced User	Man	4892	54,0
	Woman	4160	46,0
	Total	9052	100,0

When we look at both advanced and basic users in Table 3, male users are in the majority, albeit with a slight difference. The mean age was 39.82 years for group 1 and 34.68 years for group 2. This situation shows that people in the middle and elderly classes are more basic users.

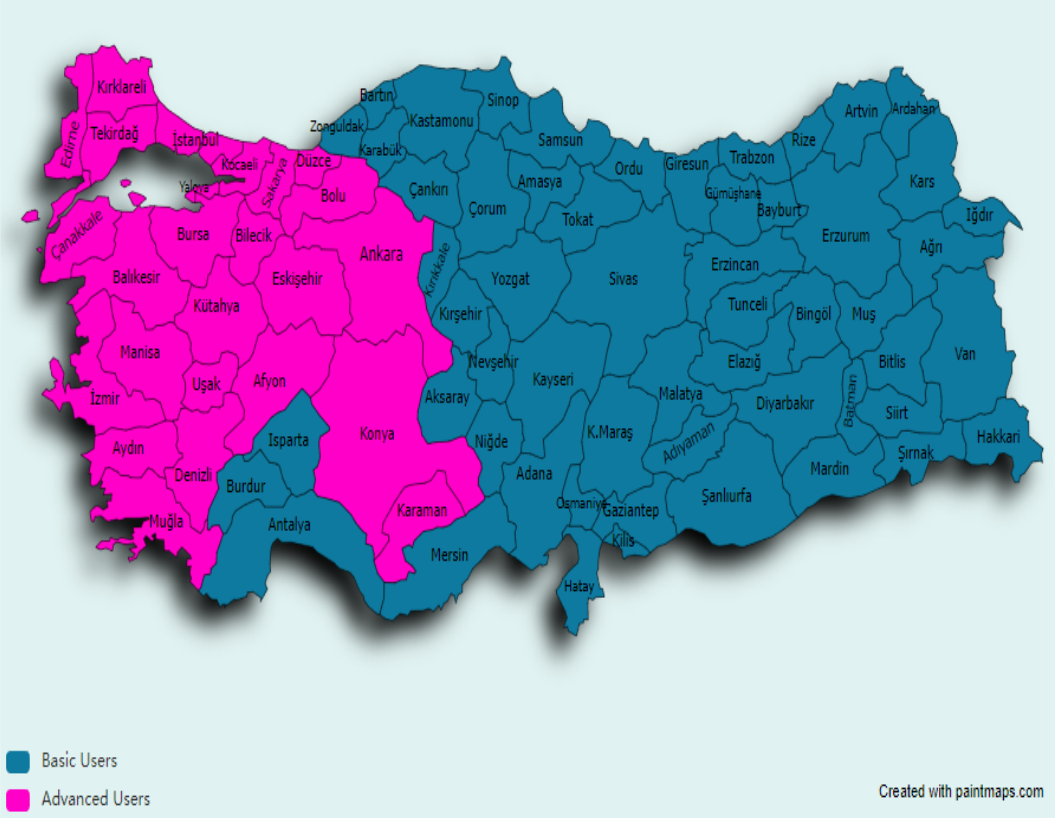
Table 4. Clustering Rates by Regions

Using Level	**Region**	**Freq.**	**Percent**	**Region**	**Freq.**	**Percent**
Basic Users	Istanbul	1869	12,2	Central Anatolia	1001	6,6
	West Marmara	869	5,7	Western Black Sea	989	6,5
	Aegean	1530	10,0	Eastern Black Sea	1000	6,5
	East Marmara	1338	8,8	Northeast Anatolia	800	5,2
	Western Anatolia	1383	9,1	Middle East Anatolia	1001	6,6
	Mediterranean	1614	10,6	Southeast Anatolia	1882	12,3
	Total	**15276**	**100,0**			
Advanced Users	Istanbul	1706	18,8	Central Anatolia	526	5,8
	West Marmara	552	6,1	Western Black Sea	613	6,8
	Aegean	1136	12,5	Eastern Black Sea	380	4,2
	East Marmara	1011	11,2	Northeast Anatolia	269	3,0

	Western Anatolia	1107	12,2	Middle East Anatolia	412	4,6
	Mediterranean	887	9,8	Southeast Anatolia	453	5,0
	Total	**9052**	**100,0**			

The distribution of data into advanced user and basic user clusters is illustrated in Table 4. The advanced usage of information technology is notably higher in the Istanbul region, whereas Northeast Anatolia has the lowest advanced user usage rate. In terms of basic user usage, Istanbul ranks second. While advanced user and basic user usage rates in most regions are similar, Southeast Anatolia exhibits a relatively high advanced-basic user difference of 7.3%. On the other hand, Northeast Anatolia had the lowest rates of advanced users and basic users' ICT use. In general, the regions of Central Anatolia, Western Black Sea, Eastern Black Sea, Northeast Anatolia, Middle East Anatolia, Southeast Anatolia, and the Mediterranean have higher rates of basic users. On the other hand, Istanbul, West Marmara, Aegean, East Marmara, and Western Anatolia regions have higher rates of advanced users. The map of Turkey formed according to the results obtained is shown in the figure below.

Figure 1. The map of Türkiye in terms of user's classification



As can be seen in Figure 1, it can be said that advanced users are concentrated in the west of Turkey. Table 5 illustrates the results when comparing the

discrepancies between the two clusters following regional differences. The table examines group variations regarding usage duration, internet capabilities, e-commerce, website purchasing problems, and online information regulation behaviours.

Table 5. Comparisons of Groups by Using Features

Main Topic	Variable	t score (p-value)	Mean Difference	Std. Error Difference
Usage Duration	Frequency of Internet use	24,513 (p=0)	0,512	0,021
	Daily Internet Usage	16,884 (p=0)	0,045	0,003
Internet Capabilities	Email Operations	73,133 (p=0)	0,434	0,006
	Voice or Video Calls	16,698 (p=0)	0,066	0,004
	Sharing Social Media Content	25,209 (p=0)	0,146	0,006
	Messaging	21,071 (p=0)	0,072	0,003
	Reading News Websites	38,333 (p=0)	0,234	0,006
	Getting Health Information	39,57 (p=0)	0,235	0,006
	Getting Information About Goods or Services	51,205 (p=0)	0,319	0,006
	Getting Information About Political Issues	16,399 (p=0)	0,071	0,004
	Taking Part in Public Debate	14,042 (p=0)	0,034	0,002
	Looking for a Job	12,284 (p=0)	0,05	0,004
	Trade	25,57 (p=0)	0,104	0,004
	Internet Banking	52,216 (p=0)	0,327	0,006
	Taking Online Courses	40,612 (p=0)	0,159	0,004
	Using Educational Materials	46,636 (p=0)	0,201	0,004
E-Commerce	Buying Clothes	6,567 (p=0)	0,063	0,01
	Sports Equipment	13,705 (p=0)	0,097	0,007
	Children's toys	10,578 (p=0)	0,084	0,008

	Home Furnishings	10,755 (p=0)	0,092	0,009
	Products like CD records	4,464 (p=0)	0,014	0,003
	DVD Blu-Ray Shopping	5,931 (p=0)	0,016	0,003
	Book/Magazine	19,442 (p=0)	0,179	0,009
	Computer Tablet Phone	14,83 (p=0)	0,117	0,008
	White goods	9,663 (p=0)	0,071	0,007
	Medicines or food supplements	10,632 (p=0)	0,074	0,007
	Ordering food	23,571 (p=0)	0,233	0,01
	Eating convenience food	20,624 (p=0)	0,195	0,009
	Cosmetics	16,313 (p=0)	0,152	0,009
	Cleaning	16,417 (p=0)	0,149	0,009
	Bicycle/Motorcycle	3,788 (p=0)	0,018	0,005
	Domestic shopping	3,453 (p=0,001)	0,01	0,003
	Shopping from EU countries	2,812 (p=0,005)	0,009	0,003
	Shopping from other countries	4,312 (p=0)	0,012	0,003
Website purchasing Problems	Difficulty of use	2,076 (p=0,038)	0,012	0,006
	Warranty conditions	1,648 (p=0,099)	0,008	0,005
	Slow delivery	8,523 (p=0)	0,068	0,008
	Shipping costs	1,886 (p=0,059)	0,008	0,004
	Wrong product shipment	4,667 (p=0)	0,034	0,007
	Fraud	-0,619 (p=0,536)	-0,003	0,004
	Response to complaints	2,86 (p=0,004)	0,013	0,005

	No problems encountered	-7,315 (p=0)	-0,073	0,01
Online Information Regulation Behaviors	Reading the privacy policy	32,059 (p=0)	0,199	0,006
	Setting limits on location tracking	36,194 (p=0)	0,21	0,006
	Setting limits on social media accounts	38,245 (p=0)	0,222	0,006
	Do not allow advertising posts	38,748 (p=0)	0,239	0,006
	Prefer secure websites	42,746 (p=0)	0,249	0,006
	Request removal of information from websites	23,691 (p=0)	0,097	0,004
	Know about the use of cookies by websites for advertising purposes	50,871 (p=0)	0,305	0,006
	Use software to prevent cookies	39,044 (p=0)	0,181	0,005
	Use software to prevent activity tracking	24,982 (p=0)	0,079	0,003

Table 5 shows many differences between users with internet access only through mobile devices and individuals using other devices. At the same time, it is seen that mobile phone device users, who are defined as basic users, perform much more transactions for daily needs than the other group. This can be interpreted as people in both groups fulfilling their daily information, shopping, or other needs with mobile devices. On the other hand, the variables found not to differ between the groups were warranty conditions, shipping costs, and fraud.

CONCLUSION

This study aims to cluster individuals according to their device usage in Turkey and examine the differences between clusters. In this direction, individuals were grouped into two clusters, the first of which consisted of those with internet access only with mobile devices and the other of those who use devices such as computers and tablets in addition to mobile devices. As a result, it was observed that the groups differed on many issues. When user types are compared regionally, the advanced user rate is higher in the Marmara, Aegean, and West Anatolia regions, while the basic user rate is higher in other regions. Another finding is that

the use of mobile devices is more intense in daily life. Another noteworthy issue is that almost all participants access the internet at least once daily and carry out many transactions such as social media, e-commerce, and banking. Therefore, especially individuals should pay attention to the following issues in terms of internet usage;

- Making optimum use of the ease and efficiency of access to the Internet,

- Avoiding addictive internet use,

- The pressure on social interaction is manageable,

- Information overload: the ability to distinguish between a lot of necessary and unnecessary information, true or false,

- Attention to security and confidentiality issues,

- Paying attention to situations such as time spent on the internet, sleep irregularity, standing still for a long time, staying still,

- Balancing online and offline life,

- To get the maximum benefit from the Internet, it is necessary to be digitally literate.

In conclusion, while the internet offers numerous advantages for everyday possibilities, individuals must focus on their online habits and maintain a healthy balance between virtual and real-world activities. This includes being aware of potential downsides and mitigating them for overall well-being.

SUGGESTIONS

As a managerial professional, it is recommended that investments be focused on the Northeast Anatolia region, as current studies indicate a lower usage of information technologies in this area. Furthermore, it is essential to note that the utilization of information technologies tends to decrease after age 40. As such, campaigns tailored to users over 40 may effectively increase usage. Researchers should consider augmenting future studies by increasing the number of years being compared, altering the study methodology or the number of clusters, or obtaining more comprehensive data by increasing the number of variables evaluated.

ACKNOWLEDGMENTS

The authors would like to thank the Turkey Statistical Institute for the data. The views and opinions expressed in this manuscript are those of the authors only and do not necessarily represent the views, official policy or position of the Turkey Statistical Institute.

REFERENCES

Baym, N. K. (2010). Personal connections in the digital age. John Wiley & Sons.

Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future

research. Educational Technology Research and Development, 55(3), 223-252.

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), 20160115.

Rogers, E. M. (1962). Diffusion of innovations. Free Press.

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). Free Press.

Vorderer, P., Klimmt, C., & Ritterfeld, U. (2012). Enjoyment: At the heart of media entertainment. Media Psychology, 14(4), 303-310.

Warschauer, M. (2003). Technology and social inclusion: Rethinking the digital divide. The MIT Press.

Berg, M. (2022). Digital technography: A methodology for interrogating emerging digital technologies and their futures. Qualitative Inquiry, 28(7), 827-836.

Perera, R. H. A. T., & Abeysekera, N. (2022). Factors affecting learners' perception of e-learning during the COVID-19 pandemic. Asian association of open universities journal, 17(1), 84-100.

Russell, G., Finger, G., & Russell, N. (2000). Information technology skills of Australian teachers: Implications for teacher education. Journal of Information Technology for Teacher Education, 9(2), 149-166.

Hollman, A. K., Hollman, T. J., Shimerdla, F., Bice, M. R., & Adkins, M. (2019). Information technology pathways in education: Interventions with middle school students. Computers & Education, 135, 49-60.

Marler, J. H., Liang, X., & Dulebohn, J. H. (2006). Training and effective employee information technology use. Journal of Management, 32(5), 721-743.

Buhalis, D. (1998). Strategic use of information technologies in the tourism industry. Tourism Management, 19(5), 409-421.

Drias, H., Cherif, N. F., & Kechid, A. (2016). K-MM: A hybrid clustering algorithm based on k-means and k-medoids. In Advances in Nature and Biologically Inspired Computing: Proceedings of the 7th World Congress on Nature and Biologically Inspired Computing (NaBIC2015) in Pietermaritzburg, South Africa, held December 01-03, 2015 (pp. 37-48). Springer International Publishing.

Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image segmentation using K-means clustering algorithm and subtractive advanced user clustering algorithm. Procedia Computer Science, 54, 764-771.

Kuo, R. J., Ho, L. M., & Hu, C. M. (2002). Integration of self-organizing feature map and K-means algorithm for market segmentation. Computers & Operations Research, 29(11), 1475-1493.

Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2006, March). Medical image segmentation using k-means clustering and improved

watershed algorithm. In 2006 IEEE southwest symposium on Image Analysis and interpretation (pp. 61–65). IEEE.

Teknomo, K. (2006). K-means clustering tutorial. Medicine, 100(4), 3.