




Paweł Kaczmarczyk 

The State University of Applied Sciences in Plock, Faculty of Economic Sciences and Information Technology, Department of Economics, p.kaczmarczyk@pwszplock.pl

Development of the Usage of E-services in Households, by Voivodship: a Cluster Analysis

Abstract: The aim of this work is to assess the development of voivodships in terms of the usage of e-services (e-government, e-commerce, e-health, etc.) in households, in comparison with the research on the overall ICT usage in households, in voivodships. The results of both of these studies enabled the verification of the thesis that the higher the level of overall ICT usage in households, in a given voivodship, the higher the level of e-services usage in these households. In the theoretical part of the work, the rationale for the research was presented. Therefore, the importance of the usage of e-services in households, for building an information society and consequently knowledge-based economy, has been described. The research methodology also included: linear ordering methods (Hellwig's method, methods that are non-based on the pattern of development), agglomerative hierarchical clustering method (Ward's method), and optimisation clustering method (the *k*-means method) have been discussed. The empirical part of the work involves presentation of the research results. Data from the year 2017, provided by the Central Statistical Office of Poland, was used. Within the framework of the usage of e-services in households (26 variables), the rankings of voivodships were created and the clusters of voivodships were detected. The obtained rankings and clusters of voivodships served to compare and assess voivodships in the analysed year. In the light of the results relating to the usage of the e-services, the results of ordering and clustering of voivodships within the framework of overall ICT usage in households (27 variables) were obtained. Then, the verification of the assumed thesis was conducted.

Keywords: e-services, households, voivodships, linear ordering, clustering method

JEL: C38, C43, O33

1. Introduction

Information and communication technologies (ICTs) have introduced a new paradigm in the field of business activity. This radically changed the approach to the role of ICT in social and economic development. Selected aspects of this new paradigm are as follows (United Nations Conference on Trade and Development, 2007: 1–2):

1. The use of ICTs may have a greater impact on economic development than their production. In other words, the economic impact of ICTs could be higher than their direct share in GDP as a production sector, due to the external effects caused by their use and application in various sectors of the economy.
2. One of the most important external effects is the new way of organizing production and consumption, which results in saving transaction costs and faster and better communication between companies.
3. The rapid pace of innovation growth in the ICT sector significantly reduces the costs of access to new technologies, i.e. the democratization of the use of ICT occurs.
4. ICT generates new services in the form of e-commerce, e-banking, e-government, etc., which contributes to economic efficiency.

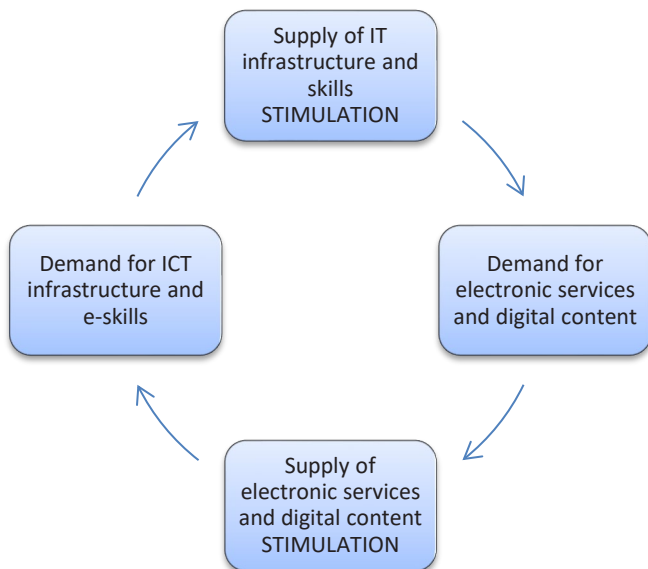


Figure 1. Areas of activities stimulating the process of computerization
Source: own elaboration on the basis of Ministerstwo Nauki i Informatyzacji, 2005

Therefore, actions stimulating the development of the information society should concern, inter alia, a development of the ICT infrastructure of the state

and in particular ensuring a universal broadband access to the Internet. These activities should also focus on the development of the system of publicly available electronic services throughout government, business and healthcare services (Figure 1).

The purpose of this study is the assessment of the development of voivodships in terms of the use of e-services (e-government, e-banking, e-insurance, e-commerce, e-education, e-culture, e-health, e-tourism) in households in comparison to the overall use of ICT in households, in individual voivodships. This issue is particularly important due to the significant impact of the development of e-services on building an information society and, consequently, a knowledge-based economy. This study poses the thesis that the higher the general use of ICT in a given voivodship, the higher the level of e-services' usage. The Central Statistical Office data for 2017 was used for the analysis.

2. E-services in the context of shaping a knowledge-based economy

E-services fall into the category of new business solutions. The new business solutions, which are introduced into the economy, are new approaches in terms of work methods, services and business models (Rust, Kannan, 2002: 3–21; Song, Korba, Yee, 2007). E-services and other new business solutions contribute to the development of the information society and digital economy (Flaga-Gieruszyńska, Gołaczyński, Szostek, 2017: 1–54). In a knowledge-based economy, the predominating services are the ones of which the market value depends to a large extent on the knowledge of service providers (managers and employees), i.e. knowledge-intensive services prevail.

There are many definitions of e-services in the subject literature, which differ in some aspects, given by such authors as e.g. Muhammad Rais and Nazariah (2003), Rowley (2006), Jeong Chun Hai (2007).

In this thesis, e-services are understood as a new formula for services' provision and therefore as satisfying consumer needs via the Internet. The virtual form of services' provision allows for a greater standardization of services and may concern the full or fragmentary servicing of e-customers as part of the service process appropriate for a given type of service (e.g. ordering a service, booking a visit, booking a ticket, borrowing a book from a library). E-services allow one to perform these activities in a more economical and effective manner. They also allow for more flexible operation of companies and their focus on creating new products or services that will generate revenue (Dąbrowska, Janoś-Kresło, Wódkowski, 2009: 9–46).

The factors determining changes in consumer behaviour towards the use of e-services include (Rust, Lemon, 2001: 85–101; Gershman, 2002: 41–47; Hofacker et al., 2007: 13–44; Dąbrowska, Janoś-Kresło, 2010: 41–65; Nof et al., 2015: 1–32):

- 1) the possibility of presenting the offer to a higher number of e-customers, to which the access by traditional methods is difficult;
- 2) a quick modification of offers and their adaptation to changing trends, prices and exchange rates, which is especially important during periods such as the financial crisis;
- 3) providing consumers with comprehensive information about services, with the possibility of using sound, animation/film or text, which contributes to the materialization of services and facilitates decision-making process;
- 4) the possibility of eliminating intermediate links within some services, which approximates the service to its recipient and may reduce the cost of the service and, consequently, the price;
- 5) the possibility of extending the working hours of a service company to 24/7 without incurring any additional costs, which, in the face of price pressure is a significant factor for the companies, and makes the company more accessible, open to the customers' needs and expectations;
- 6) a consumer seeking services on-line is anonymous, which may facilitate making independent purchase decisions;
- 7) the short time of placing an order or entering a 'purchase and sale' transaction, and in the case of some services also their consumption, cause a shortening of the time necessary to satisfy consumers' needs;
- 8) the variety of payment methods (bank transfer, payment on delivery, payment cards);
- 9) access to markets, i.e. access to both the local and international market, which is important in the era of services globalization.

E-services contribute to the development of a knowledge-based economy, because they require the transfer of knowledge in various forms: by employing appropriate people who are able to provide services, based on their knowledge; through the purchase of books, guides, magazines and other paper or electronic documents; by purchasing the relevant software; by purchasing databases and knowledge bases. E-services as knowledge-based services can be provided to: person by person, person by computer (such computer is equipped with appropriate software, having access to the right databases and knowledge bases) to a customer's computer by an employee's computer (not involving people) (Cellary, 2009: 23–47).

E-services are developing very quickly and are expected to continue to develop in the future (Sneth, Sharma, 2007: 7–12; Woodside, Quaddus, 2015: 1–4; Saprikis, 2018: 203–213). For example, the value of the e-commerce market in 2010 was equal to PLN 12.867 billion (Gemius, 2010) and in 2017 it increased to PLN 40 billion (Gemius, Izba Gospodarki Elektronicznej, 2017).

Regarding the e-services usage in households, in Poland, in comparison to other European Union countries, it is worth mentioning the Digital Economy and Society Index (DESI). This is a composite index published every year by the European Commission, measuring progress of EU countries towards a digital economy and society. It brings together a set of relevant indicators on Europe's current digital policy mix.

The above mentioned index is composed of five principal policy areas which regroup overall 31 indicators:

- 1) connectivity – fixed broadband, mobile broadband, broadband speed and prices;
- 2) human capital – basic skills and internet use, advanced skills and development;
- 3) use of internet – citizens' use of content, communication and online transactions;
- 4) integration of digital technology – business digitisation and e-commerce;
- 5) digital public services – e-government.

The position of Poland within the analysed context in comparison to other European Union countries is shown in the Figure 2.

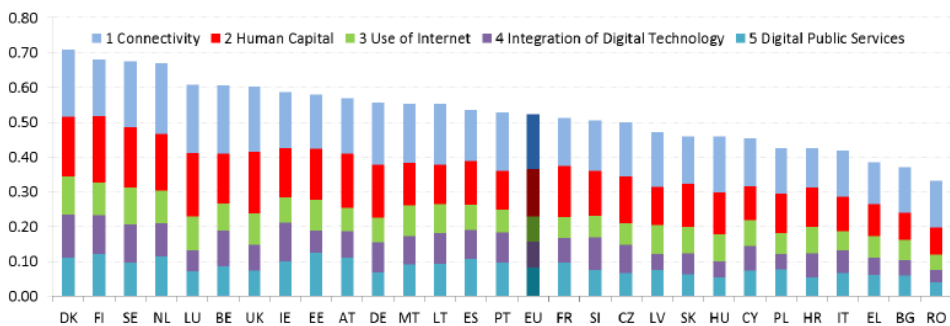


Figure 2. Digital Economy and Society Index (DESI), 2017 ranking

Source: European Commission, 2017: 1

In 2017, the DESI for Poland was 0.43 and for the EU – 0.52. The result for Poland is clearly below the European Union average. In 2017, according to this index, Poland was in the 23rd place among the EU countries. This means that Poland occupied a relatively low position in the ranking, in terms of the development of the digital economy and information society and belongs to the group of countries with poor results (European Commission, 2017: 1).

Figure 3 presents the level of use of four types of e-services (e-education, e-banking, e-government and e-commerce) in the EU countries, as well as the average level of use of these services in the EU. The presented results relate to the year 2017 and come from the Eurostat database. In the case of all analysed e-services, Poland is below the average use of these services in the European Union (Figure 3). The low positions of Poland in terms of the use of various types of e-ser-

vices is connected with the low DESI level for Poland. The empirical part of this study was devoted to the analysis of the use of e-services in individual voivodships. Among the analysed variables (in the empirical part of the paper), the variables concerning: e-health, e-commerce, e-government, e-tourism, e-banking, e-education, e-culture were presented.

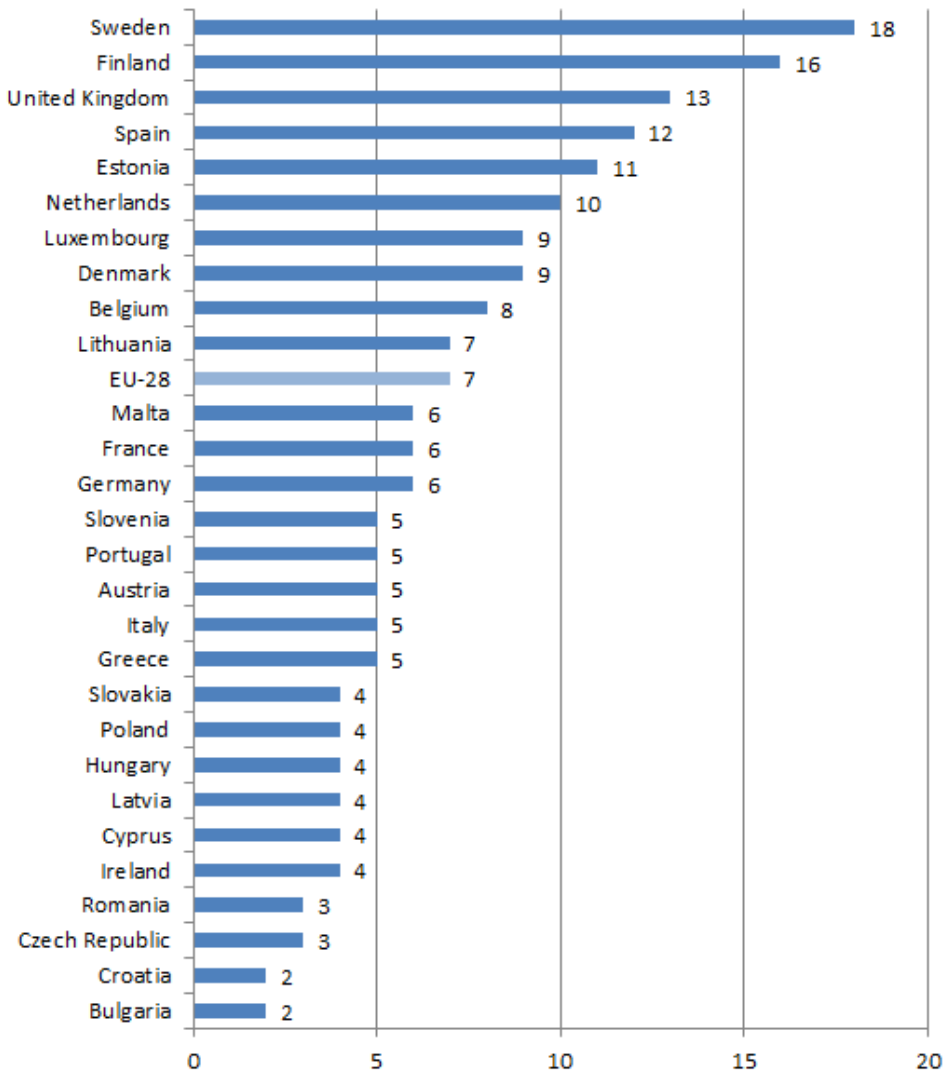


Figure 3a. The use of selected e-services in the EU countries, in 2017 (% of individuals aged 16 to 74).
Individuals using the internet for doing an online course

Source: own elaboration on the basis of Eurostat database

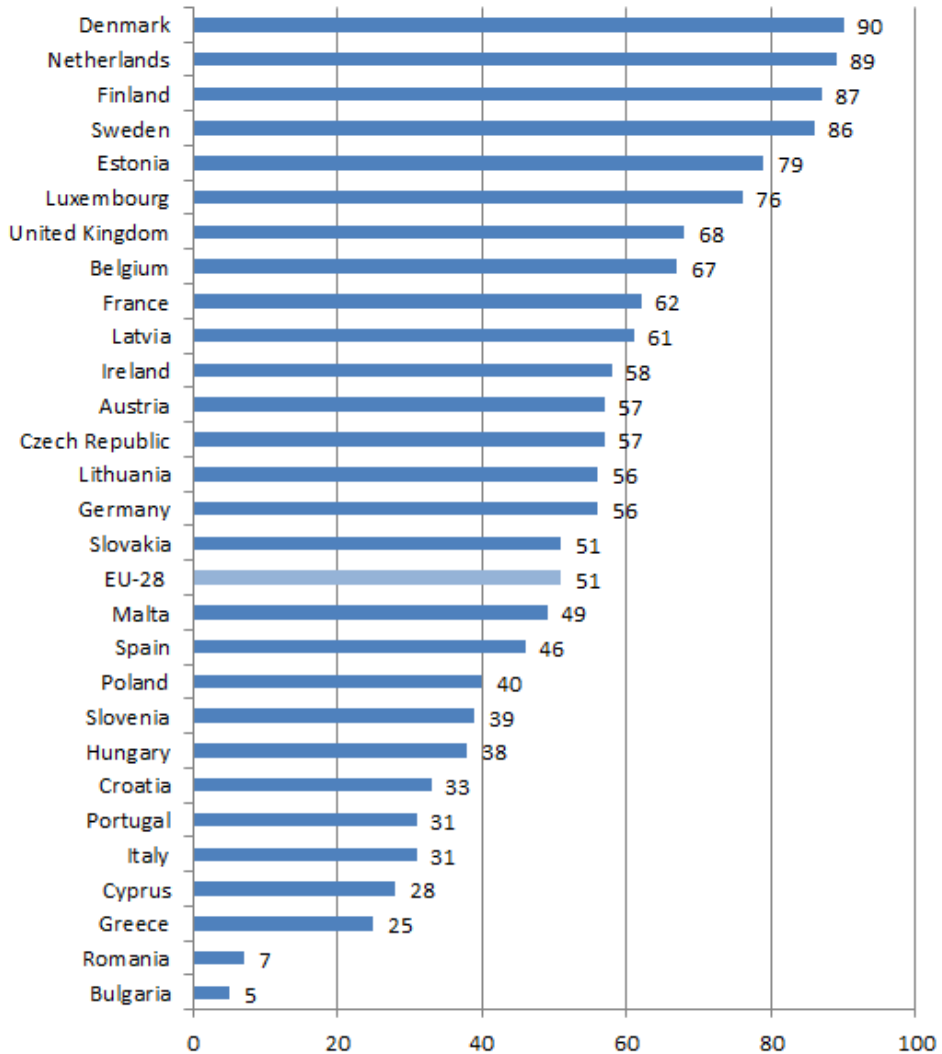


Figure 3b. The use of selected e-services in the EU countries, in 2017 (% of individuals aged 16 to 74).

Individuals using the internet for internet banking

Source: own elaboration on the basis of Eurostat database

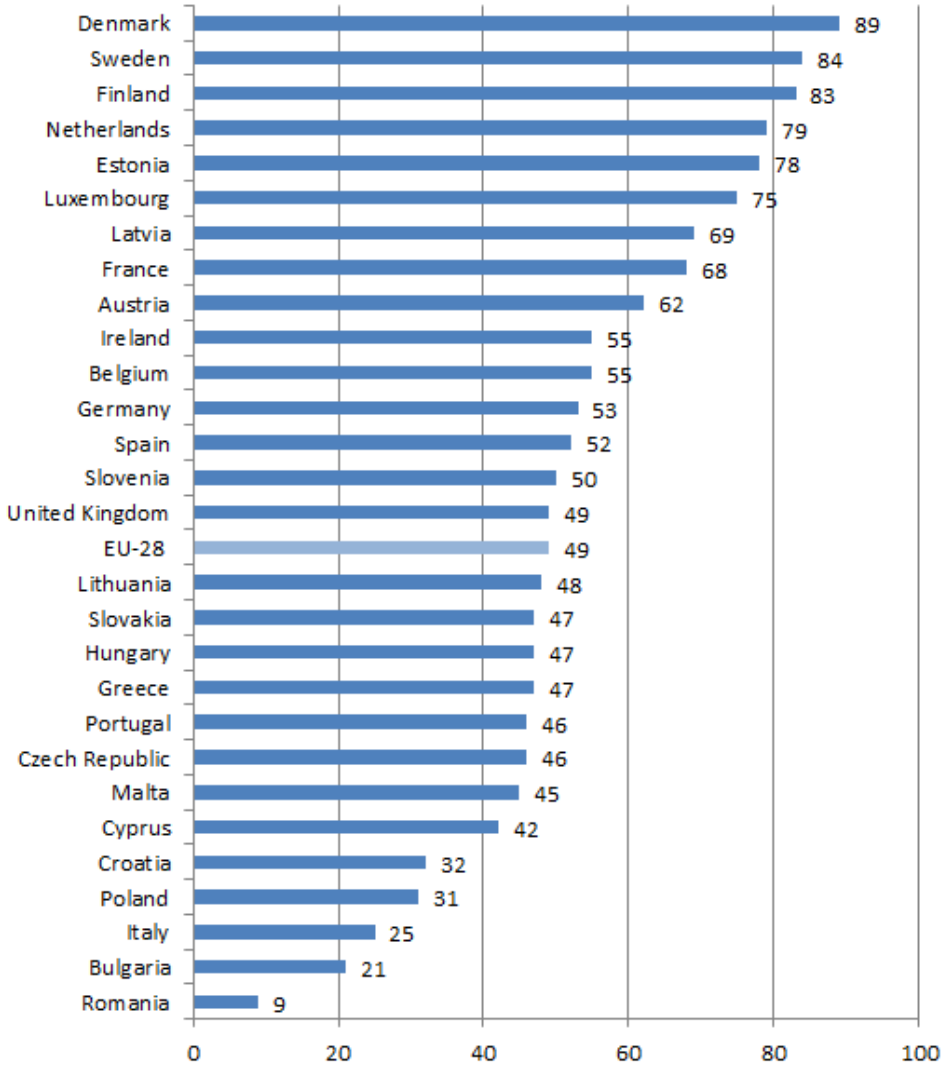


Figure 3c. The use of selected e-services in the EU countries, in 2017 (% of individuals aged 16 to 74).
Individuals using the internet for interaction with public authorities

Source: own elaboration on the basis of Eurostat database

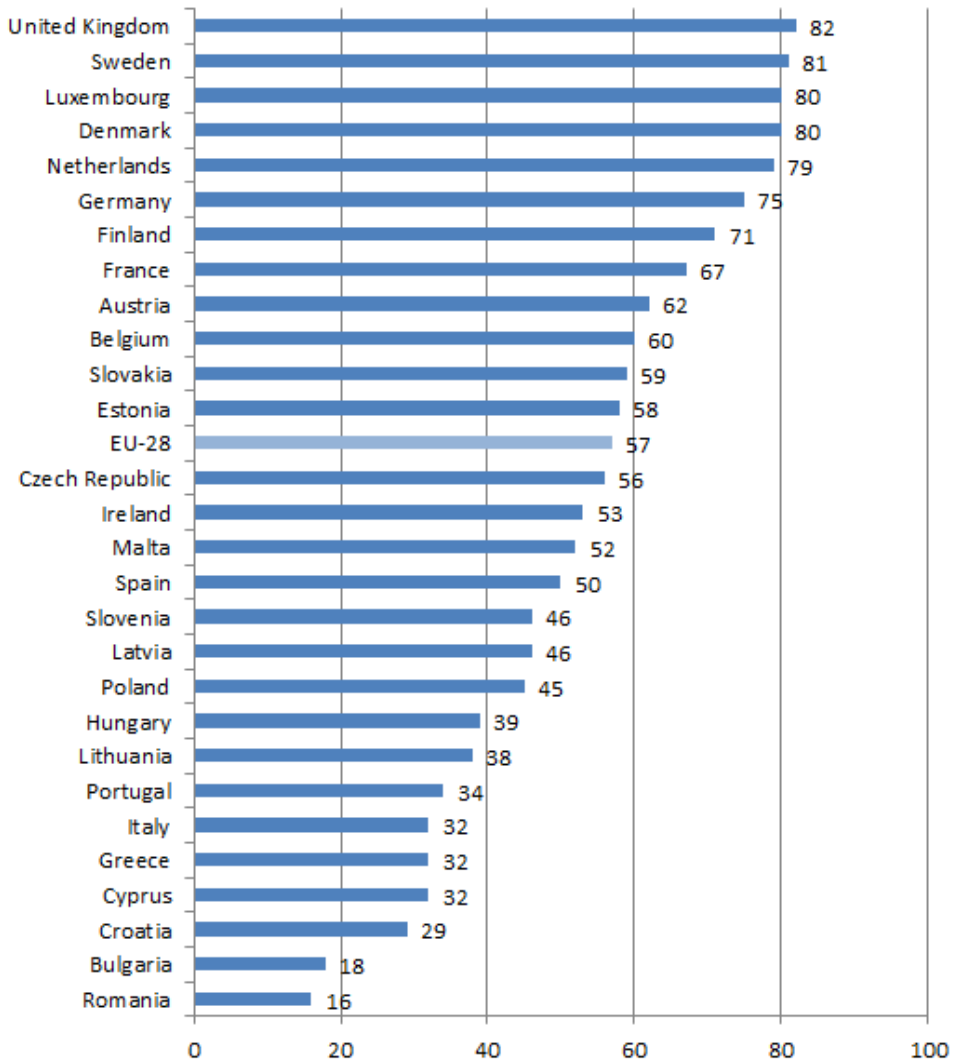


Figure 3d. The use of selected e-services in the EU countries, in 2017 (% of individuals aged 16 to 74).
Individuals using the internet for ordering goods or services

Source: own elaboration on the basis of Eurostat database

3. Applied research methodology

Empirical research began with a selection of appropriate diagnostic variables. In the process of selecting diagnostic variables, an approach based on non-statistical criteria (substantive and formal), as well as an approach based on statistical criteria were used (Panek, Zwierzchowski, 2013: 18–25). Therefore, in the first stage, a set of potential diagnostic variables was created (based on non-statistical methods). In the second stage, selected statistical procedures were applied in order to reduce the set of potential diagnostic variables and thus select the final set of diagnostic variables. This approach was used to determine the final set of variables, both in the case of analysis of the extent of the use of e-services in households, as well as in the case of analysis of the level of ICT use in households.

As part of the selection of variables according to statistical criteria, the discriminative capacity of variables as well as their information potential were verified. From the set of potential diagnostic variables, the variables for which the coefficient of variation was not greater than 0.1 were eliminated, as it is the usual threshold value. Variables that had discriminative possibilities were subjected to an information capacity study. In the information capacity test, a parametric method was used in which r^* (the critical value of the correlation coefficient) was determined based on the distribution function of Student's t -distribution. In order to increase the resistance of the results to outliers, when determining R_j (i.e. the quantities corresponding to the columns of the correlation matrix), the median was used (Młodak, 2006: 31):

$$R_j = M_j(r_{jj'}), \quad j, j' = 1, 2, \dots, m, \quad (1)$$

and then (according to the Hellwig's method) the following criteria were applied:

$$R_{j'_0} = \max_{j'} \{R_{j'}\}, \quad (2)$$

$$|r_{jj'_0}| > r^*, \quad j, j' = 1, 2, \dots, m. \quad (3)$$

The set of final diagnostic variables includes central variables and isolated variables. Thus, the satellite variables have been removed from the set of variables.

In the linear ordering procedure, when determining the aggregate measures (s_i, h_i, d_i) both the methods non-based on the pattern of development and the method based on the pattern of development were applied. In order to use methods non-based on the pattern of development, variables were subjected to unitarisation and normalisation. The values of the s_i aggregate measure were determined as arithmetic means of the unitarised features for individual objects:

$$s_i = \frac{1}{d} \sum_{j=1}^d x_{ij}, \quad (4)$$

where:

$$x_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}. \quad (5)$$

The values of the h_i aggregate measure were determined analogically, based on normalised data:

$$h_i = \frac{1}{d} \sum_{j=1}^d x_{ij}, \quad (6)$$

in which:

$$x_{ij} = \frac{x_{ij}}{\max_i x_{ij}}. \quad (7)$$

Among the methods based on the pattern of development, the Hellwig's method (1968) was used. Using this method, the values of the d_i aggregate measure were determined:

$$d_i = 1 - \frac{d_{i0}}{d_0} \quad i = 1, 2, \dots, n, \quad (8)$$

where:

$$d_{i0} = d(x_i, x_0) = \sqrt{\sum_{j=1}^d (x_{ij} - x_{0j})^2}, \quad (9)$$

$$d_0 = \bar{d}_0 + 2s_d. \quad (10)$$

The d_{i0} values were calculated based on the data for which the standardisation procedure was applied:

$$x_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}. \quad (11)$$

The Ward's method – agglomerative hierarchical clustering method was used for the process of creating voivodship clusters (Ward, 1963; Everitt et al., 2011: 77–78). In this method, the criterion of merging two clusters is an error sum of squares.

In this algorithm, the value of total within-cluster error sum of squares (E) is minimalised:

$$E = \sum_{m=1}^g E_m. \quad (12)$$

In turn, error sum of square for a given cluster is calculated by means of the following formula:

$$E_m = \sum_{l=1}^{n_m} \sum_{k=1}^{p_k} (x_{ml,k} - \bar{x}_{m,k})^2, \quad (13)$$

in which: $x_{ml,k}$ is a value of the k^{th} variable for the l^{th} object in the m^{th} cluster. Mean of the m^{th} cluster is counted in the following manner:

$$\bar{x}_{m,k} = \frac{1}{n_m} \sum_{l=1}^{n_m} x_{ml,k}. \quad (14)$$

The Ward's method was used to detect the certain number of voivodship clusters (Kopczewska, Kopczewski, Wójcik, 2016: 445). The assignment of objects to clusters was optimised using the k-means method.

The k -means method belongs to the optimisation methods of cluster analysis (McQueen, 1967). The k -means algorithm relies on iteratively updating clusters by moving each object to the group to the mean of which it was the closest. Then, the clusters' centres are calculated. The literature includes descriptions of simulation studies results aimed at comparing the effectiveness of the trace minimisation algorithms (Brusco, Steinley, 2007). As the most effective algorithms, the genetic algorithm (Maulik, Bandyopadhyay, 2000) and the variable neighbourhood search procedure (Hansen, Mladenovic, 2001) were indicated. However, the McQueen's k -means algorithm was also indicated as very effective.

In the computational part of this study, the Spearman's ranks correlation coefficient ρ and the Kendall's ranks correlation coefficient τ were applied. In order to obtain reliable results, it is worth preparing a number of rankings, based on different methods and different assumptions, and then examining them in terms of compliance of the results obtained (Kaczmarczyk, 2017). These coefficients were used to:

- 1) examine the similarity of pairs of voivodship rankings, created using various measures (s_i, h_i, d_i), within one of the two issues under consideration, i.e. the use of e-services in households or general use of ICT in households;
- 2) assess the similarity of pairs of voivodships' rankings, built with the use of the same measure (e.g. h_i), but referring to two different analysed issues.

The first approach made it possible to assess the reliability of the results obtained. The second approach allowed for the formulation of conclusions regarding the connection of the issue of general use of ICT in households with the issue of the use of e-services in households in spatial (voivodships) terms.

4. Study results

The study was carried out based on data from the year 2017. The analysed objects were voivodships V_i ($i = 1, 2, \dots, 16$). The analyses consisted of ordering and clustering these objects into two contexts: 1) the context of the level of use of e-services in households; 2) the context of the level of use of ICT in households. Table 1 presents the set of pre-selected diagnostic variables in the first of these analyses, based on substantive and formal criteria. Therefore, 26 variables were selected.

Table 1. The set of pre-selected diagnostic variables in the scope of the use of e-services in households

Variable marking	Variable name
X_1	Individuals using e-health (in %)
X_2	Individuals looking for information related to health (in %)
X_3	Individuals making an appointment to the doctor via a website (in %)
X_4	Individuals using internet banking (in %)
X_5	Individuals participating in the on-line course (in %)
X_6	Individuals using training materials other than the full on-line course (in %)
X_7	Individuals contacting the instructor/teacher through educational websites/portals (in %)
X_8	Individuals using on-line dictionaries and encyclopaedias (in %)
X_9	Individuals posting texts, photos, music and movies on websites (in %)
X_{10}	Individuals reading on-line news, newspapers or magazines (in %)
X_{11}	Individuals listening to on-line music (in %)
X_{12}	Individuals using services related to travel and accommodation services (in %)
X_{13}	Individuals using public government services via the Internet in the last 12 months (in %)
X_{14}	Individuals using the Internet to search for information on public administration websites (in %)
X_{15}	Individuals using the Internet to download official forms (in %)
X_{16}	Individuals using the Internet to send completed forms (in %)
X_{17}	Individuals using the Internet to send the tax declaration (in %)

Variable marking	Variable name
X_{18}	Individuals using the Internet to sell goods and services (in %)
X_{19}	Individuals ordering goods or services via the Internet for private use during the year (in %)
X_{20}	Individuals ordering goods or services via the Internet for private use during the last three months (in %)
X_{21}	The average value of goods and services per household, in which at least one individual purchased something in the last 3 months (in PLN)
X_{22}	The average value of goods and services per one purchasing individual (in PLN)
X_{23}	The average value of the ordered goods and services per 1 household (in PLN)
X_{24}	The average value of the ordered goods and services per 1 individual aged 16–74 in the household (in PLN)
X_{25}	Individuals ordering goods or services via the Internet during the last three months, of a value between 4,801 PLN to 12,000 PLN (in %)
X_{26}	Individuals ordering goods or services via the Internet during the last three months, of a value over 12,000 PLN (in %)

Source: own study based on GUS database

In the second analysis (Table 2), 27 diagnostic variables were pre-selected based on substantive and formal criteria.

Table 2. The set of pre-selected diagnostic variables in the scope of ICT use in households

Variable marking	Variable name
X_1	Households equipped with computers (in %)
X_2	Individuals using a computer (in %)
X_3	Individuals who use a computer regularly (in %)
X_4	Individuals using a computer in the last 3 months (in %)
X_5	Households with Internet access at home (in %)
X_6	Individuals using the Internet (in %)
X_7	Individuals who use the Internet regularly (in %)
X_8	Individuals using the Internet in the last 3 months (in %)
X_9	Individuals using mobile devices to connect to the Internet outside their home or workplace (in %)
X_{10}	The number of local hotspots installed in public places per 1000 residents (in %)
X_{11}	Households incurring ICT expenditures (in %)
X_{12}	Households incurring expenditures on computer equipment, peripheral devices, accessories and consumables (in %)
X_{13}	Households incurring software expenditures (in %)

Variable marking	Variable name
X_{14}	Average monthly ICT expenditures per household (in PLN)
X_{15}	Average monthly ICT expenditures per household incurring expenditures in this area (in PLN)
X_{16}	Average monthly expenditures on computer equipment, peripheral devices, accessories and consumables per household incurring expenditures in this area (in PLN)
X_{17}	Average monthly software expenditures per household incurring ICT expenditures (in PLN)
X_{18}	Individuals with basic information skills (in %)
X_{19}	Individuals with advanced information skills (in %)
X_{20}	Individuals with basic communication skills (in %)
X_{21}	Individuals with advanced communication skills (in %)
X_{22}	Individuals with basic problem-solving skills (in %)
X_{23}	Individuals with advanced problem-solving skills (in %)
X_{24}	Individuals with basic skills related to software (in %)
X_{25}	Individuals with advanced skills related to software (in %)
X_{26}	Individuals with basic digital skills (in %)
X_{27}	Individuals with advanced digital skills (in %)

Source: own study based on GUS database

Taking into account the statistical criterion (coefficient of variation and information capacity), both of these sets were reduced to 6 and 8 variables, respectively. Considering the significance of the dispersion of variables in the first set of pre-selected variables, the coefficient of variation turned out to be statistically significant for 23 variables: $X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{21}, X_{22}, X_{23}, X_{24}, X_{25}, X_{26}$. On the other hand, in the second set of variables, statistical significance of dispersion was found in 16 variables: $X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22}, X_{24}, X_{25}, X_{26}, X_{27}$. Further, as a result of the analysis of information capacity of variables, both reduced sets have been reduced again in the number of their elements. The results of these analyses are presented in Table 3.

Table 3. Results of verification of information potential of the diagnostic variable

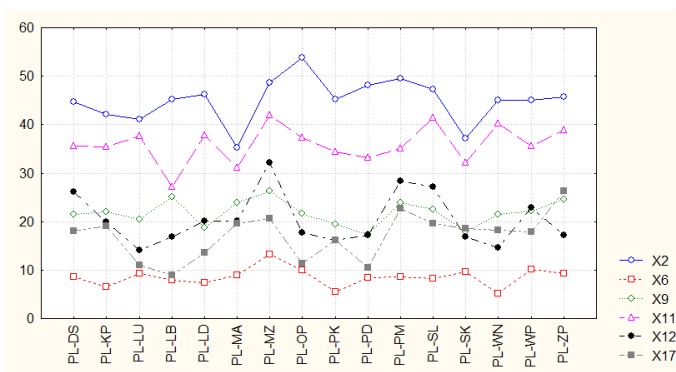
Central or isolated variable	Satellite variables
Results in terms of e-services usage in households (23 variables from the first set)	
X_{12}	$X_3, X_4, X_5, X_8, X_{13}, X_{14}, X_{15}, X_{16}, X_{18}, X_{20}, X_{21}, X_{22}, X_{23}, X_{24}, X_{25}, X_{26}$
X_{11}	
X_6	X_7
X_9	

Central or isolated variable	Satellite variables
X_2	
X_{17}	
Results in terms of overall ICT usage in households (16 variables from the second set)	
X_{16}	X_{14}, X_{15}, X_{27}
X_9	X_{25}
X_{13}	X_{11}, X_{12}
X_{24}	X_{22}, X_{26}
X_{18}	
X_{17}	
X_{10}	
X_{20}	

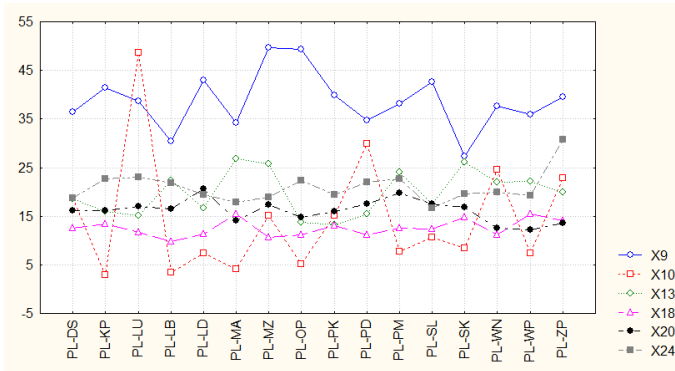
Source: own calculations

The central or isolated variables (in terms of the first phenomenon) have been treated as the final set of diagnostic variables. This set was as follows: $X_2, X_6, X_9, X_{11}, X_{12}, X_{17}$. In turn, the final set of diagnostic variables (describing the second phenomenon) has included the following central or isolated variables: $X_9, X_{10}, X_{13}, X_{16}, X_{17}, X_{18}, X_{20}, X_{24}$. The visualisation of courses of the final diagnostic variables is presented in Figure 4. In order to denote voivodships, ISO 3166–2: PL code, developed by the International Organization for Standardisation, was used.

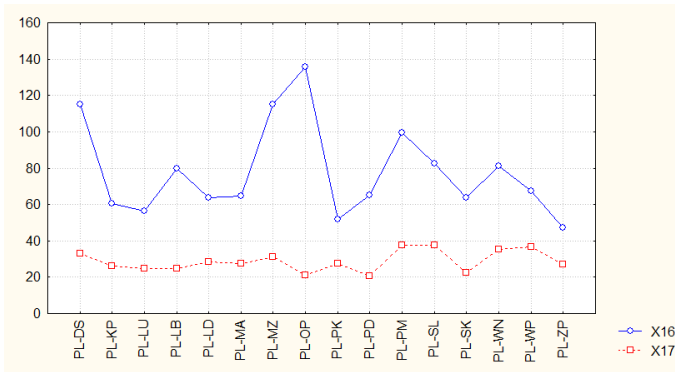
Visual analysis of Figure 4 leads to the conclusion that generally the highest values of variables are noted in the cases of the following voivodships: PL-MZ, PL-PM, PL-ZP, PL-SL. The lowest values of these variables correspond to the following voivodships: PL-PK, PL-PD, PL-LB.



The course of variables relating to the e-services usage by voivodship (in %)



The course of variables relating to the ICT usage by voivodship (in %)

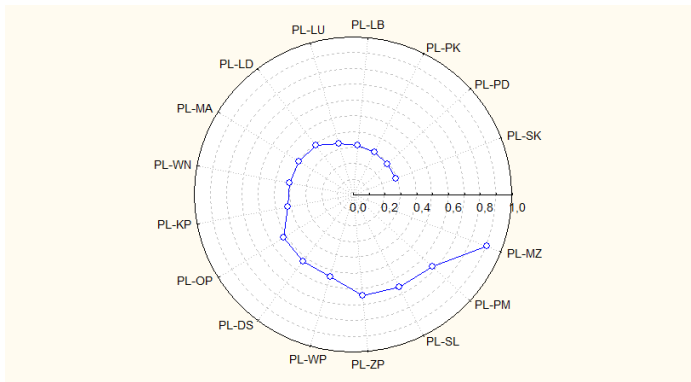


The course of variables relating to the ICT usage by voivodship (in PLN)

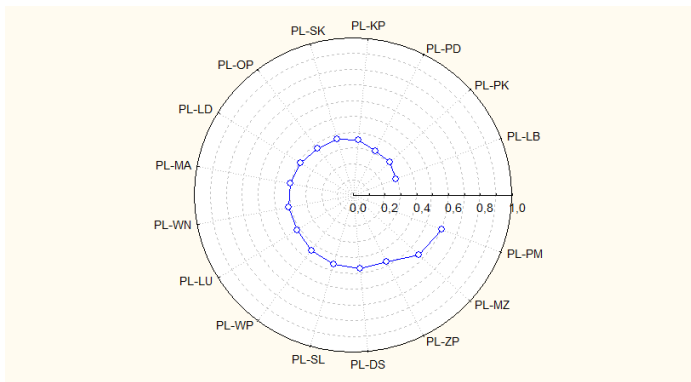
Figure 4. The course of the final selected variables

Source: own elaboration

Subsequently, rankings of voivodships have been created in two separate analyses: firstly within the framework of e-services usage in households, and secondly in terms of overall ICT usage in households. Both, the measures non-based on the pattern of development (s_i, h_i) and the measure based on the pattern of development (d_i) were used to analyse both phenomena. The obtained results using the s_i measure (i.e. the measure non-based on the pattern of development) for both studied phenomena are presented in Figure 5 as an example of the results of the rankings creation. In the further part of the results' description, it will be shown that the s_i measure is characterised by the highest level of similarity with the other measures in terms of both analysed phenomena.



Rankings of voivodships in terms of e-services usage in households



Rankings of voivodships in terms of overall ICT usage in households
 Note: The ranking was created clockwise.

Figure 5. Rankings of voivodships under the s_i aggregated measure
 Source: own calculation and elaboration

When it comes to the first analysed phenomenon (e-services usage in households) in all rankings the top three results were almost the same. This means that PL-MZ was placed in the 1st position, and the following positions were occupied by PL-PM and PL-SL. Only in one ranking (which was created under the h_i measure) the 3rd place was taken by the PL-ZP voivodship, and the PL-SL voivodship was in the 4th place. The analysis of the three last positions leads to the conclusion that in all rankings the 14th position was taken by PL-PK. The 15th and 16th places were the most frequently occupied by PL-PD and PL-LB respectively. The PL-PD voivodship was in the 15th place twice (i.e. under the s_i and h_i measure), and the PL-LB voivodship was in the 16th place twice (according to the h_i and d_i measure).

The results of correlation research with the use of Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ are juxtaposed in Table 4.

Table 4. Values of Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ in terms of e-services usage in households

ρ				τ			
Measure	s_i	h_i	d_i	Measure	s_i	h_i	d_i
s_i	1.0000*	0.9441*	0.9676*	s_i	1.0000*	0.8333*	0.8833*
h_i	0.9441*	1.0000*	0.9676*	h_i	0.8333*	1.0000*	0.8833*
d_i	0.9676*	0.9676*	1.0000*	d_i	0.8833*	0.8833*	1.0000*

Note: * denotes statistical significance at the level of 1%.

Source: own calculations

Further, the rankings were correlated from the above table (created by the use of all types of the adopted measures). In all compared pairs of rankings, values of coefficient proved to be statistically significant at the significance level of 1%, so correlation results confirm the reliability of obtained rankings.

Then, the rankings within the framework of overall ICT usage by voivodship were created with the use of the same types of measures (s_i , h_i , and d_i). Therefore, the 1st, 2nd, and 3rd places were most frequently occupied by PL-PM, PL-MZ, and PL-ZP respectively. According to only one measure (the h_i measure) the top three were the following voivodships: PL-MZ (1st position), PL-PM (2nd position), and PL-LU (3rd position). When the last three locations are taken into account, then it can be observed that the PL-LB voivodship was placed in the 16th position in all three rankings. The 15th place was most frequently taken by the PL-PK voivodship. When it comes to the 14th position, it was occupied by PL-PD, PL-KP, and PL-MA under the s_i , h_i , and d_i measure respectively.

In further part of the research, values of Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ were calculated. The received results are shown in Table 5.

Table 5. Values of Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ in terms of overall ICT usage in households

ρ				τ			
Measure	s_i	h_i	d_i	Measure	s_i	h_i	d_i
s_i	1.0000*	0.9324*	0.8206*	s_i	1.0000*	0.8000*	0.7000*
h_i	0.9324*	1.0000*	0.8529*	h_i	0.8000*	1.0000*	0.6667*
d_i	0.8206*	0.8529*	1.0000*	d_i	0.7000*	0.6667*	1.0000*

Note: * denotes statistical significance at the level of 1%.

Source: own calculations

The correlation analysis leads to the observation that all compared rankings are correlated in terms of overall ICT usage in households. All values of the correlation coefficients were statistically significant at the significance level of 1%. Therefore, the reliability of received rankings was also confirmed.

The next analysis consisted of the verification of correlation between rankings of voivodships regarding the different explored phenomena. So, in this approach to the correlation analysis, the relationship between e-services usage in households and the overall ICT usage in households can be verified. The values of the correlation analysis are shown in Table 6.

Table 6. Values of Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ between e-services usage in households and the overall ICT usage in households

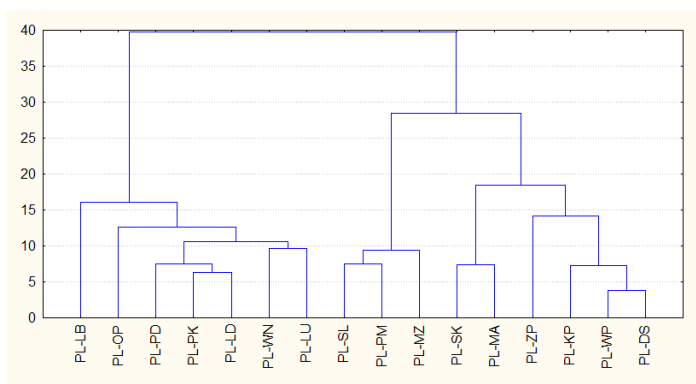
ρ between e-services and ICT				τ between e-services and ICT			
Measure	s_i	h_i	d_i	Measure	s_i	h_i	d_i
s_i	0.8412*	0.7235*	0.7206*	s_i	0.6333*	0.5667*	0.5333*
h_i	0.8706*	0.7118*	0.6647*	h_i	0.7000*	0.6000*	0.5000*
d_i	0.8529*	0.7265*	0.7471*	d_i	0.6500*	0.5833*	0.5833*

Note: * denotes statistical significance at the level of 1%.

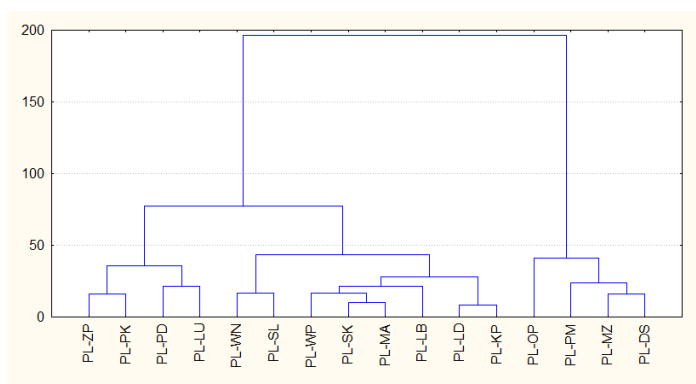
Source: own calculations

In this approach to the correlation analysis, the values of both coefficients were also statistically significant at the significance level of 1%. So, there is significant relationship between levels of development of both explored phenomena. If the development of the overall ICT usage in households is at a certain level in a particular voivodship, then the development of the usage of e-services in households is on the adequate level in this voivodship. More specifically, the higher the level of the first phenomenon in a given voivodship, the higher the level of the second phenomenon in this voivodship.

The effects of clustering the voivodships using Ward's method and the Euclidean distance for both analysed phenomena are shown in Figure 6. The analysis of both dendrograms leads to the conclusion that within the framework of the e-services usage three clusters of voivodships can be distinguished. Three clusters can also be differentiated in terms of overall ICT usage. In the case of the first phenomenon, the first (the most developed) category involves the following voivodships: PL-SL, PL-PM, PL-MZ. The second category comprises of: PL-SK, PL-MA, PL-ZP, PL-KP, PL-WP, PL-DS. Finally the third category (the least developed) consists of: PL-LB, PL-OP, PL-PD, PL-PK, PL-LD, PL-WN, PL-LU. When it comes to the second phenomenon, the first (the most developed) category contains: PL-OP, PL-PM, PL-MZ, PL-DS. The second category comprises of: PL-WN, PL-SL, PL-WP, PL-SK, PL-MA, PL-LB, PL-LD, PL-KP. The third (the least developed) category includes: PL-ZP, PL-PK, PL-PD, PL-LU.



Clustering regarding the e-services usage in households



Clustering regarding the overall ICT usage in households

Figure 6. Clustering of voivodships with the use of Ward's method

Source: own calculation and elaboration

The *k*-means method was applied to optimise the clustering of the objects. The number of clusters was set at three. In the clustering process, the distances between clusters were maximised. The effects of clustering of the voivodships using *k*-means method are shown in Figure 7.

The obtained results were very similar to the results received by means of the Ward's method. In the case of the first phenomenon, the obtained categories of voivodships differed in only two voivodships. The PL-DS voivodship was in the second category under agglomerative clustering method. This voivodship was assigned to the first category under optimization method. The PL-WN voivodship went from the third to the second category. In terms of the second phenomenon no differences were observed. The results indicate that in both of the analysed cases the categories with the same number are characterised by similar sizes. In the case of the first phenomenon, the sizes of categories (starting from the first category)

were as follows: 4, 6, 6. For the second phenomenon, the analogical categories were characterised by the following sizes: 4, 8, 4. Moreover, most voivodships constitute the categories with the same numbers for both, separately analysed phenomena.

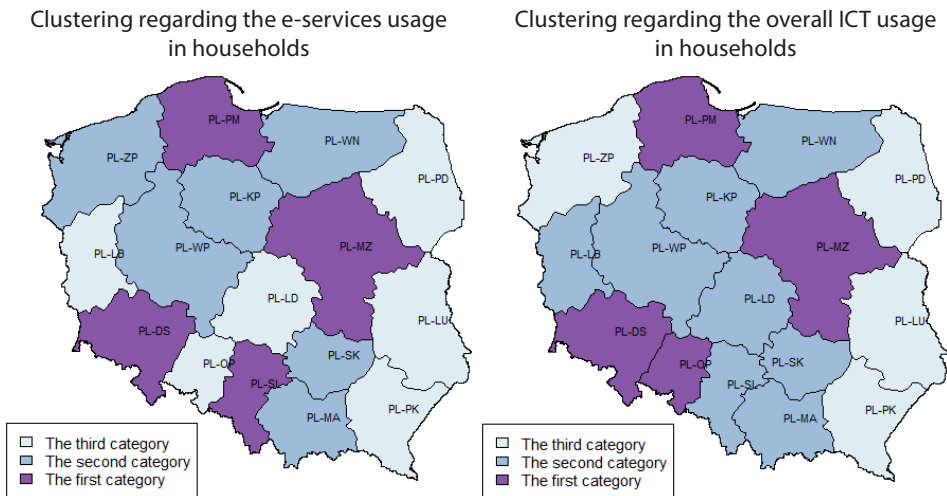


Figure 7. Clustering of voivodships by the use of *k*-means method

Source: own calculation and elaboration in the R programme

The results of *k*-means method was extended to the calculation of Euclidean distances between clusters (Table 7). Whereas Euclidean distances are presented below the main diagonals of the distance matrices for both analysed phenomena, squares of the Euclidean distances are shown above the main diagonals.

Table 7. Distances between clusters determined by means of the *k*-means method

E-services usage in households				Overall ICT usage in households			
Cluster	1	2	3	Cluster	1	2	3
1	0.0000	24.0194	38.2072	1	0.0000	269.7329	514.5231
2	4.9010	0.0000	15.7701	2	16.4235	0.0000	90.0625
3	6.1812	3.9712	0.0000	3	22.6831	9.4901	0.0000

Source: own calculations

The results of the above analysis indicate that distances between clusters in terms of the second phenomenon are greater than in the case of the first phenomenon. However, proportions of distances between clusters within both, separately analysed phenomena are similar.

5. Conclusions

The contributions of this research study consist in:

- 1) creation of rankings of voivodships in regard to the usage of e-services in households, in 2017 and then in terms of the overall ICT usage in households in the same year;
- 2) comparison and assessment of voivodships within these two frameworks;
- 3) identification and assessment of similarities between both of these analysed phenomena (with the use of three linear ordering methods and two clustering methods).

The main findings consist of the indication of the Masovian Voivodship as the leading voivodship within the scope of the e-services usage in households in 2017 and the Pomeranian Voivodship as the leading voivodship in terms of the overall ICT usage in households in the same year. Moreover, the Pomeranian Voivodship occupied the 2nd position in ranking for the e-services usage in households. In turn, Masovian Voivodship was in the 2nd place in ranking for the overall ICT usage in households. Thus, these voivodships were placed in the top three rankings in terms of both analysed phenomena. Taking into consideration the three last positions, it was discovered that the 16th place was taken by the Lubusz Voivodship in both analysed phenomena. The Subcarpathian Voivodships also took one of the last three positions in both categories (the 14th position for the e-services usage in households and the 15th position for the overall ICT usage in households). The assessment of voivodships in regard to the level of development within both of the explored phenomena was conducted on the basis of three measures (in total, 6 rankings). Thus, the reliability of the above findings was positively verified (both Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ were statistically significant at the significance level of 1% for all applied types of measures) and the result were generalised.

The obtained research results enabled the confirmation of the thesis that the higher the overall ICT usage in households in a given voivodship, the higher the level of e-services usage. The values of the Spearman's rank correlation coefficient ρ and the Kendall's rank correlation coefficient τ between two kinds of rankings (i.e. the rankings of voivodships regarding the e-services usage in households constituted the first kind of rankings and the rankings of voivodships in terms of the overall ICT usage in households constituted the second kind) proved to be significant at a very low significance level. Therefore, a higher level of e-services usage in particular voivodship is generally accompanied by a higher level of the overall ICT usage. In addition, the cluster analysis of both phenomena indicates that the majority of voivodships constitute parts of the same clusters (clusters with the same numbers).

The obtained results allowed for identification of some differences between the development of eastern and western voivodships in terms of both studied phe-

nomena. Western voivodships are generally characterised by a higher level of development in the analysed fields, i.e. they occupy higher positions in the obtained rankings and are included in the higher developed categories. However, the results of the statistical methods application do not indicate a clear differentiation of the development level between eastern and western voivodships in terms of both explored phenomena. Despite the fact that eastern voivodships are characterised by a lower level of development (ICT and e-services usage), such examples are also located in western Poland. Therefore, taking into account the concept of sustainable regional development (Ministerstwo Rozwoju Regionalnego, 2010; 2013), this finding can be treated as beneficial.

Poland takes a low position in the rankings of European Union countries in terms of e-services usage. Thus, particular indexes relating to this field in Poland (e.g. individuals using the internet to take an online course; individuals using the internet for internet banking, individuals using the internet for ordering goods or services) are meaningfully lower than average level of the same indexes for the European Union.

When it comes to the research implications, the results could be useful for central and local authorities because they indicate the particular voivodships that should increase the overall ICT usage in households and, consequently, their level of e-services usage in households. This is a very important challenge due to: the many advantages of e-services, the growing value of the e-commerce market, and the importance of e-services for strengthening a socio-economic growth and knowledge-based economy development.

References

- Brusco M.J., Steinley D. (2007), *A comparison of heuristic procedures for minimum within-cluster sums of squares partitioning*, "Psychometrika", vol. 72, no. 4, pp. 583–600.
- Cellary W. (2009), *Trzy wyzwania elektronicznej gospodarki opartej na wiedzy: biznesowe, administracyjne, społeczne*, [in:] J. Kotowicz-Jawor (ed.), *GOW – wyzwanie dla Polski*, Polskie Towarzystwo Ekonomiczne, Warszawa, pp. 23–47.
- Dąbrowska A., Janoś-Kresło M. (2010), *Konsument na rynku e-usług w krajach Europy Środkowo-Wschodniej*, Difin, Warszawa.
- Dąbrowska A., Janoś-Kresło M., Wódkowski A. (2009), *E-usługi a społeczeństwo informacyjne*, Difin, Warszawa.
- European Commission (2017), *Digital Economy and Society Index (DESI) 2017*, Brussels, http://europa.eu/rapid/press-release_MEMO-17-352_en.pdf [accessed: 26.10.2018].
- Everitt B., Landau S., Leese M., Stahl D. (2011), *Cluster analysis*, John Wiley and Sons Ltd., Chichester.
- Flaga-Gieruszyńska K., Gołaczyński J., Szostek D. (eds.) (2017), *E-obywatel, e-sprawiedliwość, e-usługi*, C.H. Beck, Warszawa.
- Gemius (2011), *E-commerce w Polsce w oczach internautów 2010*, http://pliki.gemius.pl/biuletyn_handlowy/2011_01_e_commerce_w_Polsce_internauci.pdf [accessed: 15.05.2018].

- Gemius, Izba Gospodarki Elektronicznej (2017), *E-commerce w Polsce 2017. Gemius dla e-commerce Polska*, https://branden.biz/wp-content/uploads/2018/07/E-commerce_w_Polsce_2017.pdf [accessed: 15.05.2018].
- Gershman A. V. (2002), *How Web services will redefine the service economy*, "Accenture Outlook", pp. 41–47.
- Hansen P., Mladenovic N. (2001), *J-MEANS: A new local search heuristic for minimum sum of squares clustering*, "Pattern Recognition", vol. 34, no. 2, pp. 405–413.
- Hellwig Z. (1968), *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju i strukturę wykwalifikowanych kadr*, „Przegląd Statystyczny”, vol. 15, no. 4, pp. 307–327.
- Hofacker C. F., Goldsmith R. E., Bridges E., Swilley E. (2007), *E-Services: a synthesis and research agenda*, "Journal of Value Chain Management", vol. 1, no. 1/2, pp. 13–44.
- Jeong Chun Hai I. (2007), *Fundamental of development administration*, Scholar Press, Selangor.
- Kaczmarczyk P. (2017), *Taxonomic analysis of voivodships development in terms of ICT usage in enterprises*, "Folia Oeconomica Stetinensia", vol. 17, no. 2, pp. 83–96, <https://doi.org/10.1515/fofi-2017-0020>.
- Kopczewska K., Kopczewski T., Wójcik P. (2016), *Metody ilościowe w R. Aplikacje ekonomiczne i finansowe*, CeDeWu Sp. z o.o., Warszawa.
- Maulik U., Bandyopadhyay S. (2000), *Genetic algorithm-based clustering technique*, "Pattern Recognition", vol. 33, no. 9, pp. 1455–1465.
- McQueen J. (1967), *Some methods for classification and analysis of multivariate observations*, "Computer and Chemistry", vol. 4, pp. 257–272.
- Ministerstwo Nauki i Informatyzacji (2005), *Strategia kierunkowa rozwoju informatyzacji Polski do 2013 roku oraz perspektywiczna prognoza transformacji społeczeństwa informacyjnego do roku 2020*, Warszawa, 24 czerwca.
- Ministerstwo Rozwoju Regionalnego (2010), *Krajowa strategia rozwoju regionalnego 2010–2020: regiony, miasta, obszary wiejskie*, Warszawa, <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WMP20110360423/O/M20110423.pdf> [accessed: 26.10.2018].
- Ministerstwo Rozwoju Regionalnego (2013), *Strategia rozwoju społeczno-gospodarczego Polski Wschodniej do roku 2020*, Warszawa, <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WM-P20130000641/O/M20130641.pdf> [accessed: 26.10.2018].
- Młodak A. (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa.
- Muhammad Rais A. K., Nazariah M. K. (2003), *E-government in Malaysia: improving responsiveness and capacity to serve*, Pelanduk Publications, Kuala Lumpur.
- Nof S. Y., Ceroni J., Jeong W., Moghaddam M. (2015), *Revolutionizing collaboration through e-work, e-business, and e-service*, Springer, Berlin, Heidelberg.
- Panek T., Zwierzchowski J. (2013), *Statystyczne metody wielowymiarowej analizy porównawczej. Teoria i zastosowania*, Szkoła Główna Handlowa, Warszawa.
- Rowley J. (2006), *An analysis of the e-service literature: towards a research agenda*, "Internet Research", vol. 16, no. 3, pp. 339–359, <https://doi.org/10.1108/10662240610673736>.
- Rust R. T., Kannan P. K. (2002), *The era of e-service*, [in:] R. T. Rust, P. K. Kannan (eds.), *E-service: new directions in theory and practice*, M. E. Sharpe, New York, pp. 3–21.
- Rust R. T., Lemon K. N. (2001), *E-service and the consumer*, "International Journal of Electronic Commerce", vol. 5, no. 3, pp. 85–101.
- Saprikis V. (2018), *Terrestrial vs. online gamblers: an empirical study on their perceptions towards*, [in:] M. Khosrow-Pour (ed.), *Optimizing current practices in e-services and mobile applications*, IGI Global, Hershey, pp. 203–213.
- Sneath J. N., Sharma A. (2007), *E-services: a framework for growth*, "Journal of Value Chain Management", vol. 1, no. 1/2, pp. 7–12.

- Song R., Korba L., Yee G. (2007), *Trust in e-services: technologies, practices and challenges*, Idea Group Publishing, Hershey–London.
- United Nations Conference on Trade and Development (2007), *Information Economy Report 2007–2008: Science and technology for development: the new paradigm of ICT*, New York–Geneva.
- Ward J.H. (1963), *Hierarchical grouping to optimize an objective function*, “Journal of the American Statistical Association”, vol. 58, no. 301, pp. 236–244.
- Woodside A.G., Quaddus M. (2015), *Acquiring deep knowledge of e-services adoption processes in developing nations*, [in:] M. Quaddus, A.G. Woodside (eds.), *E-services adoption: processes by firms in developing nations*, Emerald, Bingley, pp. 1–4.

Rozwój wykorzystania e-usług w gospodarstwach domowych według województw Analiza skupień

Streszczenie: Celem artykułu jest ocena rozwoju województw w zakresie wykorzystania e-usług (e-administracji, e-handlu, e-zdrowia itd.) przez gospodarstwa domowe na tle badań ogólnego wykorzystania ICT przez gospodarstwa domowe w województwach. Wyniki obu badań pozwoliły zweryfikować tezę, że im wyższy poziom ogólnego wykorzystania ICT przez gospodarstwa domowe w danym województwie, tym wyższy poziom korzystania z e-usług przez gospodarstwa domowe. W części teoretycznej pracy przedstawiono uzasadnienie badania. W związku z tym opisano znaczenie wykorzystania e-usług przez gospodarstwa domowe w budowaniu społeczeństwa informacyjnego, a w konsekwencji gospodarki opartej na wiedzy. Uwzględniono również metodologię badań. Omówiono metody porządkowania liniowego (metodę Hellwiga i metody bezwzorcowe), aglomeracyjną hierarchiczną metodę analizy skupień (metodę Warda) i optymalizacyjną metodę analizy skupień (metodę k -średnich). Empiryczna część pracy obejmuje przedstawienie wyników badań. Wykorzystano dane z roku 2017 dostarczone przez Główny Urząd Statystyczny. W ramach wykorzystania e-usług przez gospodarstwa domowe (26 zmiennych) stworzono rankingi województw i zidentyfikowano skupienia województw. Posłużyły one do porównania i oceny województw w analizowanym roku. Otrzymano również wyniki porządkowania i grupowania województw w ramach ogólnego wykorzystania ICT przez gospodarstwa domowe (27 zmiennych), jako tło wyników dotyczących korzystania z e-usług. Następnie przeprowadzono weryfikację przyjętej tezy.

Słowa kluczowe: e-usługi, gospodarstwa domowe, województwa, porządkowanie liniowe, analiza skupień

JEL: C38, C43, O33

 <p>OPEN ACCESS</p>	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (http://creativecommons.org/licenses/by/3.0/)</p> <p>Received: 2018-08-03; verified: 2018-11-22. Accepted: 2019-07-04</p>
 <p>COPE Member since 2018 JM13714</p>	<p>This journal adheres to the COPE's Core Practices https://publicationethics.org/core-practices</p>