

Paweł KACZMARCZYK¹

CLUSTER ANALYSIS OF VOIVODSHIPS IN REGARD TO SOCIAL MEDIA USAGE IN ENTERPRISES AT THE BACKGROUND OF EU STATES

The aim of this work is to assess the development of voivodships in terms of social media usage in enterprises with the use of the clustering method. In the theoretical part of the work the rationale of the research study (significance of social media in business activities of post-industrial enterprises) and research methodology were included. The empirical part of the work involves presentation of the research results. So, within the framework of social media usage by enterprises, the rankings of voivodships were created and the clusters of voivodships were detected. Data from the years 2014–2017, provided by the Central Statistical Office of Poland, was used. The results served to compare and assess voivodships in individual years of the period from 2014 to 2017. The comparisons were carried out at the background of data relating to the social media usage by enterprises in EU states in the period 2014-2017.

Keywords: social media, e-commerce, e-recruitment, voivodships, linear ordering, clustering method.

1. INTRODUCTION

In the literature of clustering of voivodships in terms of information and communication technologies (ICT), threads relating to the knowledge-based economy, information society and new managing concepts (e.g. e-business, online marketing, CRM, network management, X-engineering) can be noticed ((Papież, Śmiech [Eds.] 2016, p. 208-217; Wierzbicka, 2017; Kaczmarczyk, 2017, p. 83-96). This is due to the fact that ICT have a significant impact on the implementation and development of the above listed ideas. For this reason, ICT development is fundamental and requires conducting research. In this context, social media (characterised by many advantages) are especially important and these become more significant in socio-economic terms.

The increase in the popularity of social media is in connection with increasingly greater social media usage by enterprises in the business sphere as a new communication channel to inter alia promote their products and brands. On the other hand, social media users recommend interesting content, i.e. they share opinions about brands, products, services and advertising campaigns. Social marketing enables users to create a group of loyal customers and quickly acquire new customers. Through communication on social media, companies

¹ Paweł Kaczmarczyk, PhD, Faculty of Economic Sciences and Information Technology, The State University of Applied Sciences in Plock, 28 Głoczyńskiego Street, 09-400 Plock, Poland, p.kaczmarczyk@pwszplock.pl. ORCID: 0000-0001-5199-6328.

encourage consumers to share ideas that they can then use to work on the creation or modification of products and services. More and more often social media are useful when a company recruits employees, for example when a company seeks people with narrow specialisation. Social media has also become a tool for improving communication within the company, enabling employees to exchange opinions and ideas.

The aim of the article is to assess voivodships (regarding the development of social media usage in enterprises) on the basis of cluster analysis in each year for the period from 2014 to 2017. The social media usage is understood as the percentage of companies in the following aspects: at least one social medium is used, social networking services are used, blogs and micro-blogs are run, websites enabling making multimedia available and wiki tools are used. In order to achieve the assumed aim, the linear ordering methods (i.e. aggregated measure that was non-based on the pattern of development and aggregated measure based on the pattern of development) and Ward's method were used in the carried out analyses.

Social media can significantly support business activities of post-industrial enterprises. So, the research effects are an information basis about the possibilities of business development in individual voivodships by means of social media. In a broad context, these possibilities translates directly into socio-economic development of particular voivodships. Therefore, the results can be treated as support for the policy of digitisation and building a knowledge-based economy.

2. ROLE OF SOCIAL MEDIA IN CONTEMPORARY ENTERPRISES – THE IMPORTANCE OF RATIONALE OF THE RESEARCH

There is an extensive literature about social media in marketing (Mangold, Faulds, 2009, p. 357-365; Safko, Brake, 2009; Weber, 2009; Meerman, 2010; Castronovo, Huang, 2012, p. 117-134; Chaffey, 2015; Barger, Peltier, Schultz, 2016, p. 268-287; Van Looy, 2016). If one looks at social media as a marketing tool, then these media should be regarded severally and the communication process (which is based on social media) should be seen as an entirety. The reason is that social media are characterised by the following specific characteristics: wide (global) scope, interaction speed, a rather low cost, contribute to significant increase brand awareness, enable entrepreneurs to find suggestions and ideas, contribute to a long-term brand and product loyalty, support e-commerce and improve customer care.

From the above mentioned social media advantages, it is worth emphasising content about supporting e-commerce by social media. One can consider this support in the indirect or direct meaning. The first one is expressed as achieving high activity (positive opinion about brand and products) on websites and attracting as many fans as possible. Interesting data related to influence of social media on e-commerce are contained in the report of the Chamber of Digital Technology in Poland. According to the report, in 2016 in Poland, 61% of social media users posted opinions and information about their purchases online by means of these media. In the same year 76% of social media users stated that when they made purchasing decisions, they were guided by positive opinions of their friends. Under the mentioned data, the number of brand fans has an impact on perceiving a particular brand (i.e. whether it is trustworthy or not). Nearly 50% of social media users declared it. At the same time 60% of social media users stated that they "like" and "observe" various brands. Therefore, the above data can be treated as data related to the indirectly support of e-commerce by social media. The directly supporting e-commerce by social media is based on

enabling consumers to avoid one of processes, which is most disliked by consumers, i.e. registering or logging in to an e-shop. In 2016 in Poland 34% of consumers chose the option of logging in to other sites by means of a social media account and the majority of such sites were online shops. Currently, social media is the space in which a quite high volume of trade takes place that is indicated by the following data: 38% of social media users declared that they have already used the “buy now” button and 32% of them made a direct purchase through social media.

Summing up the content of social media impact on e-commerce, in 2016 in Poland 24% of social media users were of the opinion that a brand absent in social media is “already lost” and 21% of social media users consider that such a brand is “not trendy”. The younger they are, the more often they agree with this opinion (*Lubię to czy kupuję to?...*, 2016).

Social media are important for enterprises as they are efficient e-recruitment tool (Roe-buck, 2012; Bondarouk, Olivas-Lujan [Eds.], 2013, p. 97-120; Mulder-Williamson, Taylor, 2013; Schlesinger, 2014; Rospigliosi, Greener [Eds.] 2014, p. 57-63; Headworth, 2015; Woźniak, 2016, p. 103-124; Gravili, Fait, 2017). Companies increasingly often decide to use social media in the recruitment process, and social media popularity will also grow in the future. This is due to the fact that social media usage provides companies with the three most important long-term benefits (Gzowska, 2016, p. 79-92). First of all, social media is a significant facilitation for jobseekers because these media provide wider access and geographic scope. Regardless of the location, candidates, who are interested in a specific job offer, can apply for the offered job from any location. Secondly, social media enables recruiters to find hard-to-reach candidates. The reason for this is that the use of social media gives the opportunity to reach passive (non-seeking) candidates who are difficult to reach by other means. Thirdly, due to the application of social media, a competitive advantage can be achieved, i.e. recruiters, using social media in the recruitment process, gain an advantage over recruiters who do not use social media in this process.

However, it should be remembered that the usefulness of social media in employee recruitment depends on the type of business activities. It is worth noting that various social (business) services are not equally effective in every recruitment process. Recruiters point to the fact that the GoldenLine service is effective for seeking less qualified employees with a little professional experience or representing niche professions, as well as those whose qualifications place them at the level of not more than specialist. On LinkedIn, it is easier to conduct international recruitment processes. Here recruiters find highly qualified employees.

3. THE RESEARCH METHODOLOGY

In the initial part of the research, the rankings of voivodships (in regard to social media usage in enterprises) were created by means of methods non-based on the pattern of development (Panek, Zwierzchowski, 2013, p. 57-108). Thus, the conversion of considered variables was conducted. The conversion was carried out by the use of the unitarisation formula:

$$x_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad (1)$$

and by the use of normalisation formula:

$$x_{ij} = \frac{x_{ij}}{\max_i x_{ij}}, \quad (2)$$

in which x_{ij} is observation of i -th object and j -th variable ($i = 1, 2, \dots, n$; $j = 1, 2, \dots, d$).

The aggregated measure (the s_i measure), which bases on the unitarised variables, was computed in following manner:

$$s_i = \frac{1}{d} \sum_{j=1}^d x_{ij} \quad (3)$$

and the h_i measure (based on the normalised data) was computed as follows:

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

Within the framework of methods based on the pattern of development, Hellwig's method was used (Hellwig, 1968, p. 307-327). The aggregated measure d_i was calculated after data standardisation:

$$x_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}. \quad (5)$$

The coordinates of the pattern (ideal) object were assumed as maximal values of all variables, because all variables were stimulants. The d_i measure is determined as follows:

$$d_i = 1 - \frac{d_{i0}}{d_0} \quad i = 1, 2, \dots, n, \quad (6)$$

where:

$$d_{i0} = d(x_i, x_0) = \sqrt{\sum_{j=1}^d (x_{ij} - x_{0j})^2}, \quad (7)$$

$$d_0 = \bar{d}_0 + 2s_d. \quad (8)$$

Formula (8) means that d_0 is the sum of arithmetic mean of d_{i0} and $2s_d$ as doubled standard deviation of d_{i0} .

All above analysed aggregated measures are so constructed, that the higher value of the measure (s_i , h_i or d_i), the more advanced development of the explored phenomenon.

In order to conduct cluster analysis, Ward's method was applied (Ward, 1963, p. 236-244). This method was the analysis subject in works of many researchers (Romesburg, 2004; Gan, Ma, Wu, 2007; Walesiak, E. Gatnar [Eds.], 2004, p. 316-350). Ward's method is one of the standard agglomerative hierarchical clustering methods (including the single

linkage method; complete linkage method; group average linkage method – called also as UPGMA – the unweighted pair-group method using the average approach; centroid linkage method – known also as UPGMC – the unweighted pair-group method using the centroid approach; median linkage method – alternative name is WPGMC – the weighted pair-group method using the centroid approach). Within the framework of literature on the clustering method, comparative research, related to the correctness of discovery of given types of data structure, can be remarked. It has been proved that Ward's method (in addition to the flexible method) is the best agglomerative hierarchical method (Arabie, Hubert, Soete [Eds.] 1996, p. 341-375).

The squared Euclidean distance was used in Ward's method. This method has geometric interpretation only with assumption of this type of distance. According to Ward's method, the merging of two clusters is based on the value of sum of squared error (*SSE*). Two clusters are so joined that the increase of total within-cluster *SSE* is minimal, so *SSE* is the criterion of fusion of two clusters. *SSE* is calculated in following manner:

$$SSE = \sum_{i=1}^k SSE(C_i), \quad (9)$$

in which $SSE(C_i)$ is within-cluster *SSE* and it is determined as follows:

$$SSE(C_i) = \sum_{\mathbf{x} \in C_i} (\mathbf{x} - \mu(C_i))(\mathbf{x} - \mu(C_i))^T \quad (10)$$

or

$$SSE(C_i) = \sum_{\mathbf{x} \in C_i} \mathbf{x}\mathbf{x}^T - \frac{1}{|C_i|} \left(\sum_{\mathbf{x} \in C_i} \mathbf{x} \right) \left(\sum_{\mathbf{x} \in C_i} \mathbf{x} \right)^T = \sum_{\mathbf{x} \in C_i} \mathbf{x}\mathbf{x}^T - |C_i| \mu(C_i) \mu(C_i)^T, \quad (11)$$

where:

$$\mu(C_i) = \frac{1}{|C_i|} \sum_{\mathbf{x} \in C_i} \mathbf{x}, \quad (12)$$

$$\mathbf{x} = (x_1, x_2, \dots, x_d)^T, \quad (13)$$

so, $\mu(C_i)$ is the mean of C_i , \mathbf{x} is a vector denoting the object (point) from cluster C_i , and x_j is a scalar denoting j -th attribute of \mathbf{x} . The number of attributes d , is also called the dimensionality of the data set.

Therefore, total *SSE* is sum of individual within-cluster *SSE*. Due to the criterion, which is used in Ward's method, this method is also called as the minimum variance method.

In connection with the presented properties of Ward's method, the main characteristics of this method are juxtaposed in Table 1.

Table 1. The summary of Ward's method

Alternative name	Usually used with:	Distance between cluster are defined as:	Remarks
Minimum sum of square	Distance (requires raw data)	Increase in sum of squares within clusters, after fusion, summed over all variables	Assumes points can be represented in Euclidean space for geometrical interpretation. Tends to find same-size, spherical clusters. Sensitive to outliers.

Source: Own coverage on the basis of (Everitt, Landau, Leese, Stahl, 2011, p. 79).

If one uses the squared Euclidean distance to compute dissimilarity matrix, then the dissimilarity matrix can be analysed by means of the Lance and Williams recurrence formula (Lance, Williams, 1967, p. 373-380; Scheibler, Schneider, 1985, p. 283-304):

$$D(C_k, C_i \cup C_j) = \frac{|C_k| + |C_i|}{\sum_{ijk}} D(C_k, C_i) + \frac{|C_k| + |C_j|}{\sum_{ijk}} D(C_k, C_j) - \frac{|C_k|}{\sum_{ijk}} D(C_i, C_j), \quad (14)$$

where:

$$\sum_{ijk} = |C_k| + |C_i| + |C_j|. \quad (15)$$

If one assumes that cluster C_i and C_j will be merged into C_t cluster and this cluster will be joined with C_k cluster, then the increase in SSE can be noted in following manner:

$$\Delta SSE = \frac{|C_k| + |C_i|}{|C_k| + |C_t|} \Delta SSE_{ki} + \frac{|C_k| + |C_j|}{|C_k| + |C_t|} \Delta SSE_{kj} - \frac{|C_k|}{|C_k| + |C_t|} \Delta SSE_{ij}, \quad (16)$$

where: μ_i, μ_j – the means of clusters C_i, C_j respectively.

From the above it is concluded that an increase in total SSE is caused by an increase in SSE within the following pairs of clusters: C_k and C_i , C_k and C_j . On the other hand, the increase in SSE, that was caused by creating cluster C_t (on the basis clusters i and j) influences reduction of ΔSSE .

4. THE RESEARCH RESULTS

The research was conducted for each year from the period of four years (2014-2017). The clustered elements were voivodships V_i ($i = 1, 2, \dots, 16$). Diagnostic variables X_j , describing the clustered elements, were initially assumed on the basis of GUS data (*Spółeczeństwo informacyjne w Polsce...*, 2014; *Spółeczeństwo informacyjne w Polsce...*, 2015; *Spółeczeństwo informacyjne w Polsce...*, 2016; *Spółeczeństwo informacyjne w Polsce...*, 2017). These pre-selected diagnostic variables (in %) were following:

- X_1 – enterprises using at least one of the below mentioned social media,
- X_2 – enterprises using social networking services,
- X_3 – enterprises having blogs and micro-blogs,

X_4 – enterprises using portals enabling multimedia sharing,

X_5 – enterprises using Wiki tools.

Therefore, the above listed pre-selected diagnostic variables were relative frequencies referring to enterprises using a particular type of social media ($j = 1, 2, \dots, 5$).

In order to exam the discriminative possibilities of the pre-selected variables, the descriptive statistics were computed. The below specified descriptive statistics were computed: arithmetic mean (AM), standard deviation (SD), standard error (SE), coefficient of variation (CV) and coefficient of skewness (CS). The obtained values of the descriptive statistics enables selection of the variables having discriminative possibilities. Effects of computations of three first statistics were visualised in Figure 1.

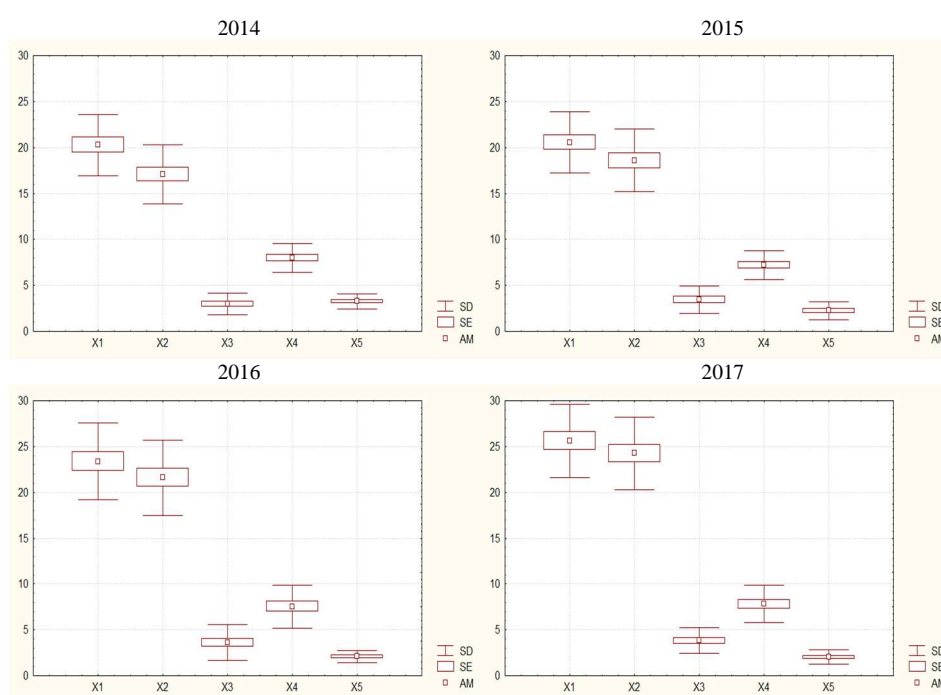


Figure 1. Box-plot on the basis of all variables in each year from the period 2014–2017

Source: Own coverage on the basis of data from GUS.

The majority of the variables, namely variables: X_1 , X_2 , X_3 are characterised by a clear growth tendency in annual AM for the period 2014–2017. In the case of variable X_5 , a decreasing tendency can be noticed in AM in the same period. It was remarked that the higher value of AM of each analysed variable is generally accompanied by higher value of SD in the period 2014–2017. In the majority of the variables (X_1 , X_2 , X_3), the highest value of AM can be observed in 2017, but in case of the three listed variables, the highest value of SD can be clear observed in 2016. Variable X_4 was characterised by the highest SD also in 2016, but the highest value of AM of this variable was in 2014. In case of variable X_5 , the highest value of SD can be observed in 2015, despite the fact that the highest value of average of

this variable was in 2014. This findings (based on the visual analysis) suggest that e.g. *CV* of X_1 , X_2 , X_3 variables in 2016 should be higher than in the next year. Except for 2014, the analysed variables were characterised by the same order in regard to value of *AM* in the assumed period. The order of the variables in this context was following: X_1 , X_2 , X_4 , X_3 , X_5 . In 2014 variable X_5 was before X_3 .

The above consideration was supplemented by computing the all previously listed descriptive statistics (*AM*, *SD*, *CV*, *CS*). The results was juxtaposed in Table 2.

Table 2. Values of the descriptive statistics

Var.	<i>AM</i>	<i>SD</i>	<i>CV</i>	<i>CS</i>	<i>AM</i>	<i>SD</i>	<i>CV</i>	<i>CS</i>
	2014				2015			
X_1	20.2813	3.3373	0.1646	0.8007	20.5500	3.3435	0.1627	0.5937
X_2	17.0875	3.2133	0.1880	1.1316	18.6000	3.4215	0.1840	0.5305
X_3	2.9875	1.1831	0.3960	-0.3555	3.4438	1.4656	0.4256	1.3560
X_4	7.9938	1.5373	0.1923	0.0943	7.1875	1.5671	0.2180	0.7278
X_5	3.2688	0.8268	0.2529	-0.5075	2.2375	0.9722	0.4345	0.9055
	2016				2017			
X_1	23.3813	4.2010	0.1797	1.2595	25.6188	4.0298	0.1573	-0.5770
X_2	21.6063	4.0977	0.1897	1.5271	24.2500	3.9623	0.1634	-0.6657
X_3	3.6063	1.9292	0.5350	2.1145	3.8250	1.4224	0.3719	0.8476
X_4	7.5500	2.3571	0.3122	1.1954	7.8250	2.0260	0.2589	0.2681
X_5	2.0875	0.6752	0.3234	0.9082	2.0375	0.7788	0.3822	1.1733

Source: Own calculation on the basis of data from GUS.

In individual years for the period 2014-2017, values of *CV* (in cases of the all variables) were higher than the assumed threshold 10%. It means that a discriminative possibility of the all variables sufficed and the all variables in all years from the period of mentioned four years should be regarded in the clustering process. When considering the analysis of relative dispersion, it can be remarked that variable X_3 is characterised by the highest level of *CV* in the great majority of analysed years (i.e. three years). On the other hand, variable X_1 is characterised by the lowest value of *CV* in all analysed years. Within the framework of the analysis of *CS*, one notices mostly right-handed asymmetry. In cases of this asymmetry, there are more values lower than *AM* in comparison to values higher than *AM*.

Then the courses of the variables were analysed and compared. In Figure 2, values of the all variables in each year for the period 2014-2017 were presented. ISO 3166-2:PL code (implemented by the International Organization for Standardization), was applied to denote voivodships. Each panel of Figure 2 related to another variable from the set of the analysed variables (X_1 , X_2 , X_3 , X_4 , X_5).

The visual analysis leads to the remark that within the framework of variable X_1 , the Masovian Voivodship was placed on the 1st position in each year from the period 2014-2017. Furthermore, within the top three (in terms of variable X_1) the following voivodships were placed most frequently: Lower Silesian (once the 2nd position and twice the 3rd position), Lesser Poland (once the 2nd position and twice the 3rd position). When it comes to the last three positions (14th, 15th, 16th), Holy Cross Voivodship can be first of all distinguished. This voivodship was twice 3rd and twice 2nd. Within the last three voivodships, Wamian-Masurian Voivodship and Subcarpathian Voivodship occurred also occupying them frequently often (each of them thrice).

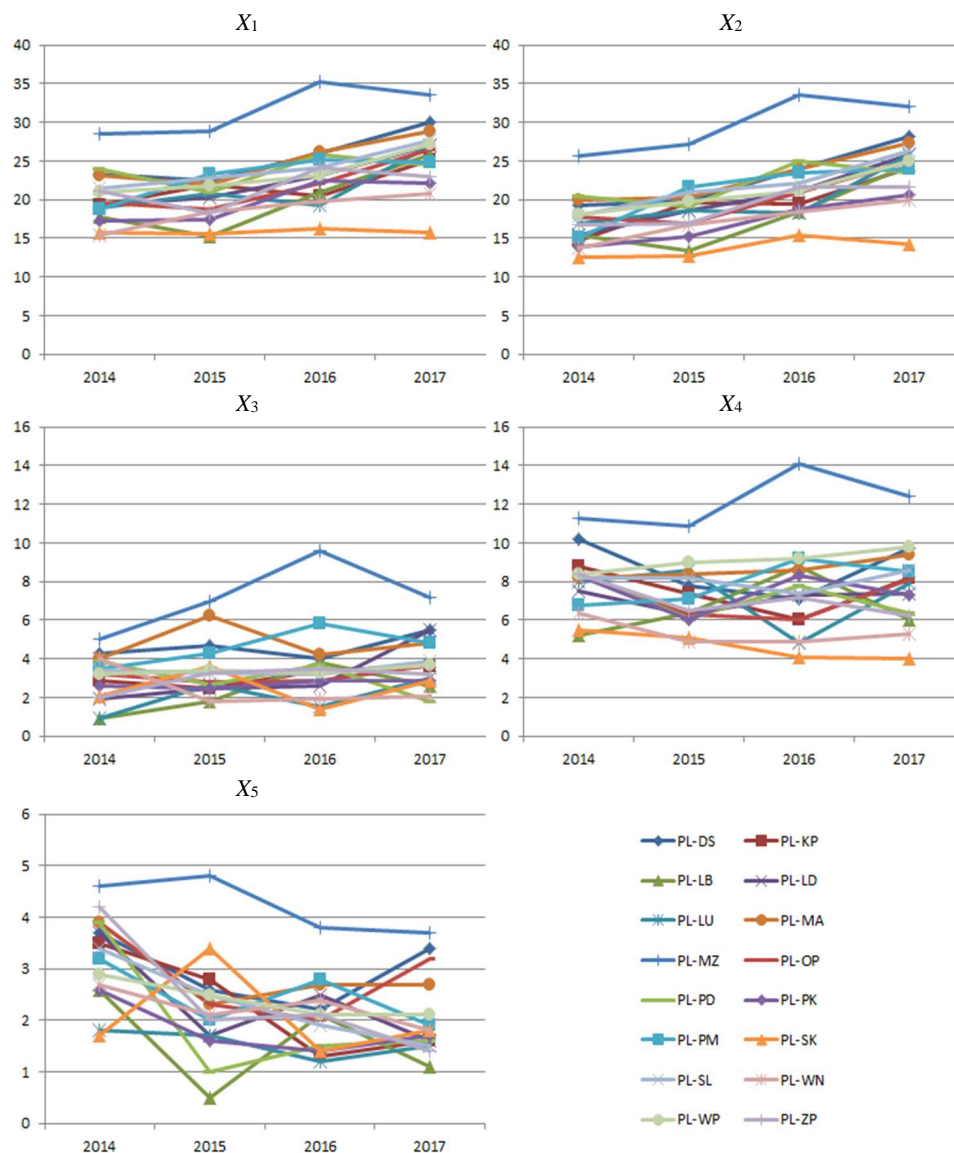


Figure 2. The courses of analysed variables (in %) in the period 2014–2017

Source: Own coverage on the basis of data from GUS.

Much similarity to the above remarks can be noticed within the variable X_2 . Masovian Voivodship was the leader in each year. Lesser Poland Voivodship was thrice in the top three. Lower Silesian Voivodship was twice in locations of this range. Referring to the three last places, the last position was taken by Holy Cross Voivodship in all the analysed years. Besides (within these places), the most frequently occurring voivodships (and their frequency) were the same as in the case of variable X_1 .

In terms of variable X_3 , one can remark that Masovian Voivodship was 1st in each year from analysed period. Within the 2nd and 3rd position, Lower Silesian and Lesser Poland Voivodship are characterised by the most frequency of occurrence. In turn, Wamian-Masurian, Lubusz, and Lublin Voivodship were the most frequently placed among three last voivodships.

With reference to voivodships positions in the field of variable X_4 , Masovian Voivodship was also placed on the 1st position in each year for the period of the analysed four years. Within the two consecutive positions, Greater Poland and Lower Silesian Voivodship were present thrice and twice respectively. It is worth noting that Holy Cross, Wamian-Masurian, and Lubusz Voivodship were usually present in the 14th, 15th and 16th places.

Considering variable X_5 , Masovian Voivodship was placed in the 1st position (in all the cases). Lesser Poland Voivodship was twice in the 3rd position. The same results were observed in the case of Opole Voivodship. In turn, Lubusz, Lublin and Subcarpathian Voivodship were placed the most frequently (each of them thrice) in the three last voivodships.

Summarising the above considered results, Masovian Voivodship was the absolute leader in all considered aspects. It can be also mentioned that Lower Silesian and Lesser Poland Voivodship were especially often seen in the top three in the vast majority of the analysed variables (four variables). If the three last locations are taken into account, then it can be emphasised that Warmian-Masurian, Holy Cross and Subcarpathian Voivodship were placed the most frequently in cases of majority of analysed variables. In this sense, Warmian-Masurian Voivodship was listed in cases of four variables in this range of locations. In the cases of Holy Cross and Subcarpathian Voivodship, it can be noticed that each of them was listed as the most frequent in terms of three variables.

In order to verify the above mentioned remarks, the rankings for particular years were created. The three aggregated measures were used. The selected results of correlation analysis between rankings were presented in Table 3.

Table 3. Values of Kendall's rank correlation coefficient τ

τ							
2014				2015			
Measure	s_i	h_i	d_i	Measure	s_i	h_i	d_i
s_i	1.0000*	0.9500*	0.9500*	s_i	1.0000*	0.9167*	0.9500*
h_i	0.9500*	1.0000*	0.9333*	h_i	0.9167*	1.0000*	0.8667*
d_i	0.9500*	0.9333*	1.0000*	d_i	0.9500*	0.8667*	1.0000*
2016				2017			
s_i	1.0000*	0.9667*	0.9667*	s_i	1.0000*	0.9500*	0.9667*
h_i	0.9667*	1.0000*	0.9333*	h_i	0.9500*	1.0000*	0.9833*
d_i	0.9667*	0.9333*	1.0000*	d_i	0.9667*	0.9833*	1.0000*

Note: * denotes statistical significance at 1% level

Source: Own calculation on the basis of data from GUS.

Due to the fact that several rankings were prepared (based on the different aggregated measures), their conformity was examined. This approach makes it possible to assess reliability of obtained results and conclusions. Therefore, values of Spearman's rank correlation coefficient ρ and Kendall's rank correlation coefficient τ were calculated. The calculations referred to correlation in particular pairs of various measures (i.e. between: s_i and h_i , s_i and d_i , h_i and d_i) within the framework of individual year in the explored period. All

obtained values of both coefficients indicate that correlation in the all compared pairs of rankings is statistically significant ($p < 0.01$). Presentation of the results (in Table 3) was limited to the values of Kendals's rank correlation coefficient τ .

The rankings under the h_i measure were presented as an exemplary in Figure 3.

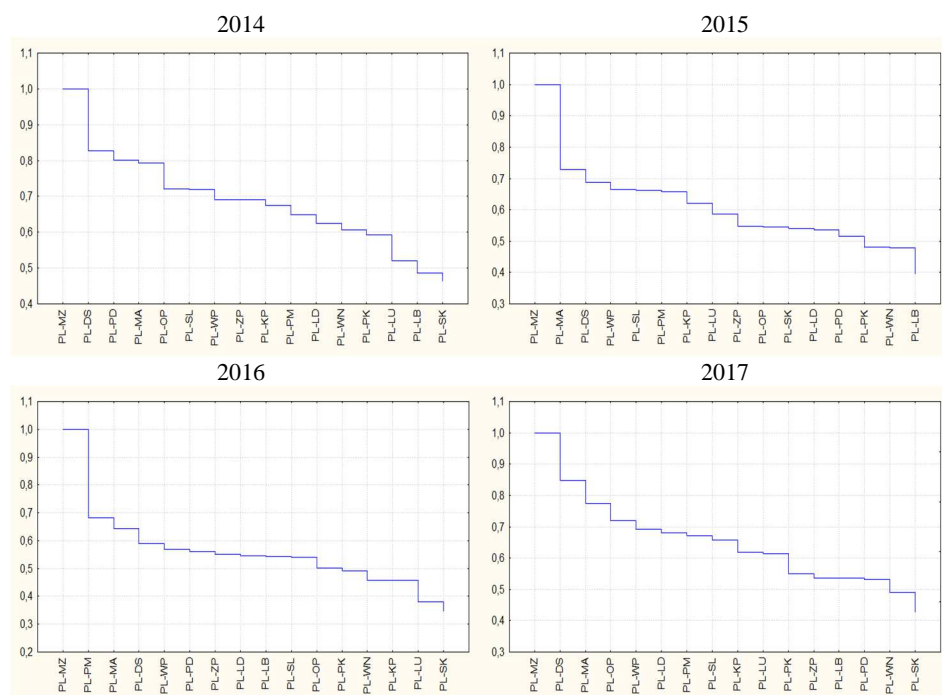


Figure 3. Rankings of voivodships under the h_i measure

Source: Own coverage on the basis of data from GUS.

When it comes to the top three, it was noticed that the results (obtained by the use of all types of applied measures) were the same within the framework of particular year. Masovian Voivodship was placed in the 1st position in all analysed years. The following two positions were the most frequently occupied by Lower Silesian and Lesser Poland Voivodship. In 2014 and 2017 Lower Silesian Voivodship was in the 2nd place and in 2015 this voivodship was 3rd. Lesser Poland Voivodship was once located in the 2nd position and then twice in the 3rd place. It was observed in 2015, 2016 and 2017 respectively. So, the result related to the top three was coincident with the result of visual analysis of particular variables in the period of four years. The analysis of the three last positions leads to the statement that the lowest level of explored phenomenon was observed in Holy Cross Voivodship (thrice in the last position under the all measures). In 2015 Lubusz Voivodship was located in the last position. This voivodship was also in the 15th position in 2014. In case of this voivodship, the results was confirmed by the all measures. Warmian Masurian Voivodship was also located in the range of three last places twice. It means, in the 15th position

in 2015 (under the measure s_i and h_i) and in the 15th position in 2017 (under the three measures). The 14th position (in 2014 under the measure s_i and h_i) and the 15th position (in 2016 under the three measures) were taken by Lublin Voivodship. The result was very similar to the remarks from visual analysis. The frequency of occurrence in the first three places according to the carried out visual analysis (i.e. the analysis of the course of individual variables in time) does not necessarily coincide directly with places in the ranking because in the linear ordering of objects, the distances between objects are the most important.

The effects of clustering the voivodships by means of Ward's method and the square of the Euclidean distance for four examined years are presented in Figure 4. On the other hand, the spatial visualisation of the clusters of voivodships was also presented in the next figure (Figure 5).

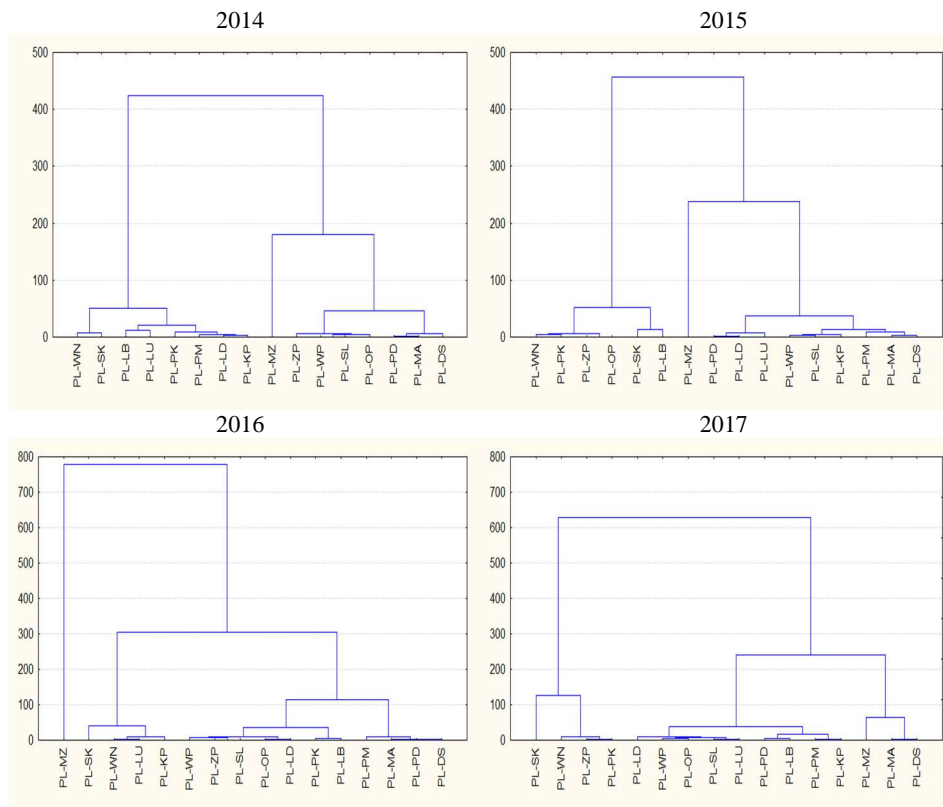


Figure 4. Clustering of voivodships by means of Ward's method

Source: Own calculations on the basis of data from GUS.

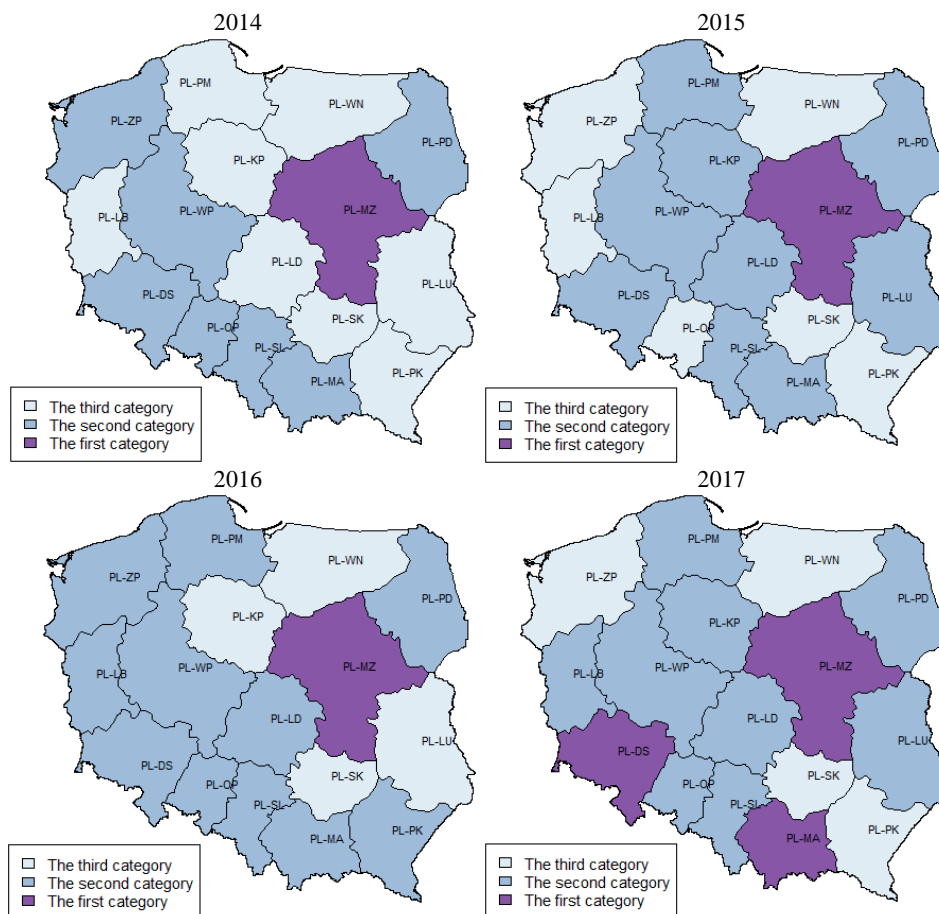


Figure 5. Spatial visualization of the voivodships clusters fixed by the use of Ward's method

Source: Own coverage (on the basis of dendrograms analysis) in R programme.

The analysis of the dendrogram for 2014 leads to the conclusion that three categories of voivodships can be differentiated. The first (the best) category comprises only one voivodship, i.e. Masovian Voivodship. The second category consists of the seven following voivodships: West Pomeranian, Greater Poland, Silesian, Opole, Podlachian, Lesser Poland, and Lower Silesian. The third (the least developed) category contains: Holy Cross, Warmian-Masurian, Lubusz, Lublin, Subcarpathian, Pomeranian, Lodz, and Kuyavian-Pomeranian Voivodship. Therefore, this cluster involved eight objects. In comparison to the previous year, in 2015 the third category was reduced to six voivodships. Namely, Lublin, Pomeranian, Kuyavian-Pomeranian and Lodz Voivodship went to the second cluster, but West Pomeranian and Opole tumbled to the third category. So, the number of objects in the second category increased to nine. In 2016 West Pomeranian and Opole Voivodship returned to the second category. It caused the third category to consist of only four

voivodships, whereas the second cluster involved eleven voivodships. Important changes were found in 2017. The first category was extended by Lesser Poland and Lower Silesian Voivodships, which went from the second category. After changes in terms of second and third category, the second category included nine objects and the third category was comprised of four objects.

To summarise, the number of objects in the first category and second category increased in the period 2014–2017. In the first case, the number increased from one to three, and in the second category – from seven to nine (in 2016 the number even equalled 11). If the changes of number of third category are reflected, then one can describe them as decreasing. The number of objects in this category decreased from nine to four. In connection with this, the detected changes in analysed structure of voivodships can be assessed as positive.

The maps are so created that the better the category of a voivodship, the darker is its colour (Figure 5). Therefore, the figure enables assessment of spatial dispersion. Especially important is the identification of a differentiation between the east and the west of Poland. The maps do not point out a clear dispersion between eastern and western voivodships in terms of analysed categories. The greatest differentiation in the analysed period was in 2014. The majority of western voivodships were in the second category, whereas the majority of eastern voivodships were in the third category. In the subsequent years, this structure of voivodships was blurred. In this context, the detected clusters and their changes in the analysed period can therefore be said to be beneficial. What is more, it was perceived that the clusters, which were distinguished in terms of overall ICT usage by enterprises (Kaczmarczyk, 2017, p. 83-96), coincided with the clusters presented in Figure 5. In order to compare the results, the effect of clustering voivodships from quoted work was set in Figure 6.

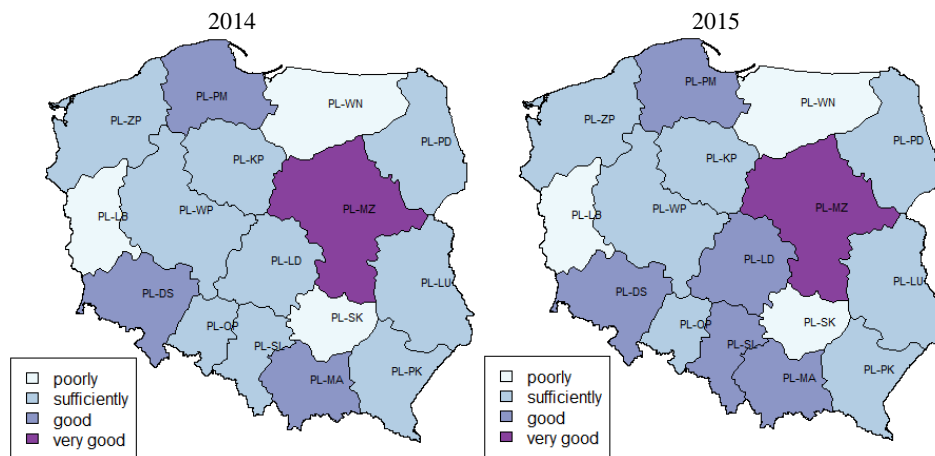


Figure 6. Clustering of voivodships under the k -means method

Source: (Kaczmarczyk, 2017, p. 93).

In the stated work, seventeen variables, associated with ICT usage in enterprises, were used to detect clusters of voivodships. Therefore, the following variables (being relative frequencies) were inter alia used in the clustering of voivodships: enterprises using computers, enterprises with access to the Internet, enterprises with broadband access to the Internet, employees using computers in enterprises, employees using computers with access to the Internet, enterprises placing orders by means of computer networks, enterprises receiving orders by means of computer networks, enterprises running big data analyses. In quoted research, four clusters (very good, good, sufficiently, and poor) were discovered for the years 2014 and 2015 separately. There are evident analogies in the compared works. The cluster, which is characterised by the highest level of ICT usage (the cluster termed as very good), contained one voivodship, namely Masovian Voivodship. Holy Cross, Warmian-Masurian, and Lubusz Voivodship were classified into the lowest cluster (identified as poor), so it is coincident with the findings assembled in Figure 5. Therefore, statement can be formulated that a voivodship belonging to a cluster characterised by a particular level of social media usage in enterprises depends on its belonging to a cluster referring to overall ICT usage in enterprises.

Enterprises (in %) using any social media in the EU states are shown in Figure 7.

It is worth emphasising that in Poland the level of social media usage was the lowest in the whole of the European Union in 2015-2017. The percentage of enterprises using any social media was assumed as the indicator in the below described analysis. The data was extracted from Eurostat database. Then, the data was ordered by the decreasing value of the indicator.

Therefore, Poland has occupied the last position in the period 2014-2017 and the results for Poland increased from 22% in 2014 to 27% in 2017. Meanwhile, the values of the indicator for the European Union (28) were 36%, 39%, 45%, 47% for 2014, 2015, 2016, and 2017 respectively.

Therefore, in 2014 the result for Poland was 14 percentage points lower than the average for the European Union (28). Then (in 2015) Poland was 17 percentage points lower than the European Union (28). In 2016 and 2017 the distance amounted 20 percentage points. It can be detected that the distance between Poland and the European Union has stabilised in the last analysed year. It means in 2017 the distance did not increase. In 2017 the difference between Poland and the European Union was at the unchanged level.

Despite stopping the distance increase, this analysis indicates a serious challenge for the Polish policy of digitalisation within the framework of the considered issue. Moreover, the results of Malta (the leader in the all analysed years) increased from 66 to 74 per cent for individual years from the period 2014-2017. So, taking into account the results of Malta, the indicator for Poland was 44, 50, 46, and 47 percentage points lower in 2014, 2015, 2016, and 2017 respectively. In comparison to the percentage of enterprises using any social media in Malta, results of Poland are very low. The percentage of enterprises using social media (percentage of enterprises regardless of kind of social media – social networking services, blogs and micro-blogs, portals enabling multimedia sharing, Wiki tools) in Masovian Voivodship were 28.5%, 28.9%, 35.3%, and 33.5% in 2014, 2015, 2016, and 2017 respectively. Therefore, the results of even the leading-edge voivodship in this context, namely Masovian Voivodship, are clearly lower than in case of average for the European Union (28).

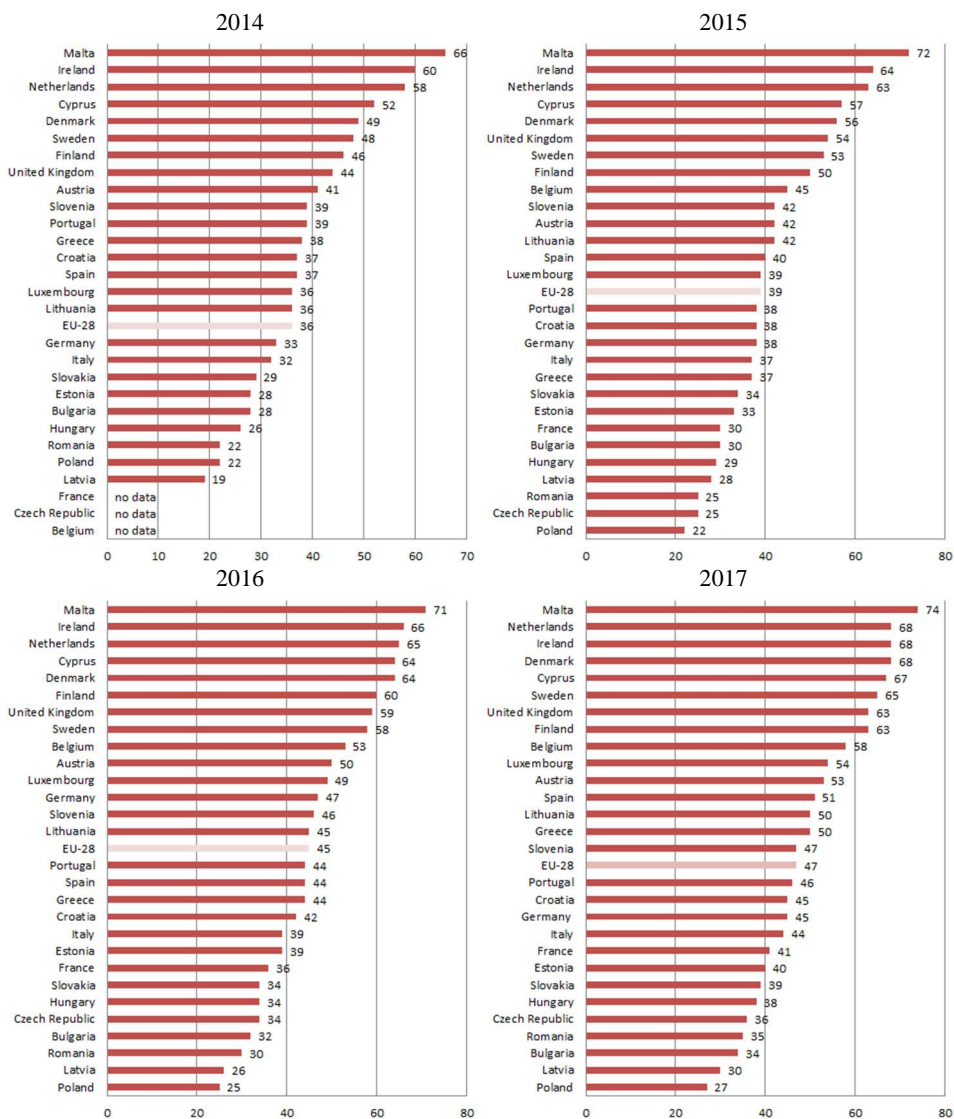


Figure 7. Enterprises using any social media in the EU states (in %)

Source: Own coverage on the basis of Eurostat database.

5. CONCLUSIONS AND RECOMMENDATIONS

With regard to social media usage by enterprises, only Masovian Voivodship existed in the first (the best category) from 2014 to 2016. So, the first category consisted originally of one object but in 2017 the number of voivodships in this category was extended to three. It was caused by inclusion of two additional following voivodships: Lower Silesian and Lesser Poland. In the analysed period of four years, the second category was also extended

from seven to nine objects. On the other hand, the third category was reduced from nine to four voivodships. Warmian-Masurian and Holy Cross Voivodship were located in the third category in consecutive years (without exception) for the period 2014-2017. The above mentioned voivodships (in the first and third cluster) were also located very similarly in the vast majority of analysed variables in individual years for the period 2014-2017. In conclusion, the changes of the structure of voivodships in the analysed field are beneficial. However, from the above is concluded a recommendation for regional policy connected with digitalisation and building a knowledge-based economy. So, the voivodships belonging especially to the third category should escalate social media usage in enterprises, which is obviously connected with overall ICT usage by enterprises in these voivodships. It is especially important from the point of view of the results of the EU states in terms of social media usage in enterprises, because Poland was located in the last position (taking into account the percentage of enterprises using any social media).

In turn, the analysis of spatial dispersion leads also to the positive conclusion. The maps do not indicate a clear dispersion between eastern and western voivodships in terms of analysed categories. It can be perceived that the dispersion between eastern and western regions decreased in the analysed period. The majority of western voivodships was in the second category in individual years for the period of 2014-2017, but some eastern voivodships were also classified into the second category. Podlachian Voivodship was in the all analysed years in the second category and Lublin Voivodship was in the second category in 2015 and 2017. Notwithstanding, the eastern voivodships with the lowest social media usage in enterprises should concentrate on intensification of the explored phenomenon.

REFERENCES

- Barger, V., Peltier, J.W., Schultz, E. (2016). *Social media and consumer engagement: a review and research agenda*. "Journal of Research in Interactive Marketing", 10/4, p. 268–287.
- Bosseli, R., Cesarini, M., Mercurio, F., Mezzanzanica, M. (2014). *How the social media contributes to the recruitment process?* [in:] Rospigliosi A., Greener S., Eds., *The proceedings of the European Conference on Social Media*, Academic Conferences and Publishing International Limited, UK, p. 57–63.
- Castronovo, C., Huang, L. (2012). *Social media in an alternative marketing communication model*. "Journal of Marketing Development and Competitiveness", 6/1, p. 117–134.
- Chaffey, D. (2015). *Digital business and e-commerce management. Strategy, implementation and practice*. Pearson: Harlow.
- Everitt, B., Landau, S., Leese, M., Stahl, D. (2011). *Cluster analysis*, Chichester: John Wiley and Sons, Ltd.
- Gan, G., Ma, C., Wu, J. (2007). *Data clustering: theory, algorithms, and applications*, American Statistical Association and the Society for Industrial and Applied Mathematics, Philadelphia, PA, Alexandria, VA.
- Girard, A., Fallery, B., Rodhain, F. (2013). *Integration of social media in recruitment: A Delphi study* [in:] Bondarouk T., Olivas-Lujan M.R., Eds., *Social media in human resources management*. Bingley: Emerald Group Publishing Limited, p. 97–120.
- Gravili, G., Fait, M. (2017). *Social recruitment in HRM: a theoretical approach and empirical analysis*. Bingley: Emerald Group Publishing Limited.
- Spółeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2010–2014*, GUS, Warszawa 2014 [access: 15.03.2018]. Access on the internet: <http://www.gus.gov.pl>.

- Spółeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2011–2015*, GUS, Warszawa 2015 [access: 15.03.2018]. Access on the internet: <http://www.gus.gov.pl>.
- Spółeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2012–2016*, GUS, Warszawa 2016 [access: 15.03.2018]. Access on the internet: <http://www.gus.gov.pl>.
- Spółeczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2013–2017*, GUS, Warszawa 2017 [access: 15.03.2018]. Access on the internet: <http://www.gus.gov.pl>.
- Gzowska, W. (2016). Wybrane media społecznościowe jako narzędzie wspierające rekrutację. „Zarządzanie Zasobami Ludzkimi”, 5, p. 79–92.
- Headworth, A. (2015). *Social media recruitment: how to successfully integrate social media into recruitment strategy*, London and Philadelphia: Kogan Page.
- Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju i strukturę wykwalifikowanych kadr. „Przegląd Statystyczny”, 4, p. 307–327.
- Lubię to czy kupuję to? Jak media społecznościowe wspierają sprzedaż? Raport Izby Gospodarki Elektronicznej, 2016 [access: 15.03.2018]. Access on the internet: <http://www.ecommerce-polska.pl/files/3414/7137/0127/LubieToCzyKupujeToSocialCommerceSierpień2016.pdf>.
- Kaczmarczyk, P. (2017). Taxonomic analysis of voivodships development in terms of ICT usage in enterprises. „Folia Oeconomica Stetinensia”, 17/2, p. 83–96. DOI: 10.1515/foli-2017-0020.
- Lance, G.N., Williams, W.T. (1967). A general theory of classificatory sorting strategies: 1. Hierarchical systems. „Computer Journal”, 9/4, p. 373–380.
- Mangold, W.G., Faulds, D.J. (2009). Social media: The new hybrid element of the promotion mix. „Business Horizons”, 52/4, p. 357–365. DOI: 10.1016/j.bushor.2009.03.002.
- Meerman, S. (2010). *The new rules of marketing an PR: How to use news releases, blogs, podcasting, viral marketing and online media to reach buyers directly*. New Jersey: John Wiley & Sons Inc.
- Milligan, G. (1996). Clustering validation: results and implications for applied analyses [in:] Arabie P., Hubert L.J., De Soete G., Eds., *Clustering and classification*. Singapore: World Scientific, p. 341–375.
- Mulder-Williamson, K., Taylor, E.M. (2013). *Simplify social media for recruiting: a step-by-step handbook for implementing social media*. Bloomington: iUniverse Inc..
- Panek, T., Zwierzchowski, J. (2013). Statystyczne metody wielowymiarowej analizy porównawczej. Teoria i zastosowania. Warszawa: SGH, p. 57–108.
- Roebuck, K. (2012). *Recruiting on social media with LinkedIn, Facebook and Twitter: high-impact strategies: what you need to know: definitions, adoptions, impact, benefits, maturity, vendors*. Lightning Source, La Vergne, TN.
- Romesburg, H.C. (2004). *Cluster analysis for researchers*. North Carolina: Lulu Press, 2004.
- Safko, L., Brake, D.K. (2009). *The social media bible. Tactics, tools & strategies for business success*. New Jersey: John Wiley & Sons Inc..
- Scheibler, D., Schneider, W. (1985). Monte Carlo test of the accuracy of cluster analysis algorithms: a comparison of hierarchical and nonhierarchical methods. „Multivariate Behavioral Research”, 20/3, p. 283–304.
- Schlesinger, R. (2014). *The suitability of social media for headhunters to recruit managers from and for the fashion business*. Hamburg: Anchor Academic Publishing.
- Strożek, P., Jewczak M. (2016). Information and communication technologies in Poland – regional perspective [in:] Papież M., Śmiech S., Eds., *The 10th Professor Aleksander Zelias*

- International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference proceedings*. Cracow: Foundation of the Cracow University of Economics, p. 208–217.
- Van Looy, A. (2016). *Social media management: technologies and strategies for creating business value*. Springer, Switzerland.
- Walesiak, M. (2004). *Metody klasyfikacji* [in:] Walesiak M., Gatnar E., Eds., *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*. Wrocław: Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, p. 316–350.
- Ward, J.H. (1963). *Hierarchical grouping to optimize an objective function*. “*Journal of the American Statistical Association*”, 58/301, p. 236–244.
- Weber, L. (2009). *Marketing to the social web. How digital customer communities build your business*. New Jersey: John Wiley & Sons Inc.
- Wierzbicka, W. (2017). *Regional Differentiation of Information Infrastructure in Poland in the Context of Building a Knowledge-Based Economy*, Institute of Economic Research Working Papers, 141 [access: 15.03.2018]. Access on the internet: http://www.badania-gospodarcze.pl/images/Working_Papers/2017_No_141.pdf.
- Woźniak, J. (2016). *Social media as an e-recruitment tool for different generations: methodological considerations and pilot study*. “*Human Resource Management*”, 6, p. 103–124.

DOI: 10.7862/rz.2019.mmr.3

The text was submitted to the editorial office: January 2019.

The text was accepted for publication: March 2019.

