Multivariate statistical analysis of information society in Poland

Justyna Brzezińska[1]

Abstract

Widespread access to high-speed Internet, user-friendly public e-services, increasing digital competence of the society are main goals for the following years due to the latest reports published by the Central Statistical Office of Poland. They are included, inter alia, in the Operational Programme Digital Poland. This technological development is also connected with the development of economic areas and public services. The rapidly increase of significance of information and electronic services, and thus, of the application of information and communication technologies (ICT), in surrounding economy, public administration (central and local government) and in the everyday life of citizens has triggered a new transformation trend – a transformation towards the information society. This term describes a society for which the processing of information with the use of ICT solutions creates significant economic, social and cultural value.

In this paper we present the current state, main aspects, vision and mission of the information society in Poland, as well as statistical analysis of information society in Poland using multivariate statistical methods. All calculations based on data from Central Statistical Office are conducted in R software.

Keywords: *information society in Poland, multivariate statistical analysis, categorical data analysis, R software*
JEL Classification: O35, C30, C39
AMS Classification: 62P20, 52H25

1 Introduction

The Information Technology (IT) has a great impact on people, society, and economy. We present briefly some facts on information society in 2015 in Poland that will help to understand the speed and impact of Information Technology on economy (Kondratiuk-Nierodzińska 2016, Piech 2007). Information society and problems related to information society in modern word was also mentioned by Stephanidis (1998), Britz (2008), Sourbati (2011), Pohle (2015). From 2012 the percentage of large companies with access to the Internet was close to 100%, which indicates saturation phenomena in this group entities. In 2015 access to the Internet had 92.7% of the companies using mostly broadband (91.9%). Mobile broadband was used by over 61.5% of companies. The rate of enterprises with their own website in 2015 amounted to 65.4%. Nearly two-thirds of companies used their own website to present catalogs of products and services they serve. In 2014, every fifth company

[1] University of Economics in Katowice, Faculty of Management, Department of Economic and Financial Analysis, 1 Maja 50, 40-287 Katowice, e-mail: justyna.brzezinska@ue.katowice.pl.

consisted of orders via computer networks, and every tenth received orders via the Internet .In 2015 almost half of large enterprises used the social media . The services of cloud computing in 2015 has used one quarter of large enterprises (source: www.gus.pl).

The financial perspective for the period 2014 to 2020 has opened up new possibilities for supporting the development of the information society. Widespread and easy access to the Internet, user-friendly public e-services, increasing digital competence of the society are aims for the following years. Information and communication technologies (ICT) have bigger and bigger influence on our daily life. In particular, the Internet is creating a new layer of participation of individuals in the societal and economic life.

In this paper we present the comprehensive description of the dynamic change of the information society' in Poland, as well as multivariate statistical analysis of the strengths and weaknesses of the information society's situation in Poland using up-to-date factual data. We also present the application of advanced statistical R software with the use of multivariate statistical analysis methods for the comparative analysis of information society in Poland.

2 Facts and figures on the ICT sector and products

Based on reports published by the Central Statistical Office (www.stat.gov.pl) in the year 2014 the number of enterprises hiring 10 or more persons in the ICT sector amounted to 2146 (6.3% increase in comparison to the previous year) among which 89.1% offered ICT services. Almost three quarters of ICT service enterprises provided IT services. In comparison with 2011 the number of ICT enterprises was increasing systematically and was higher by 24.5% (of which service enterprises by 29.2%). The number of persons employed in the sector amounted to 196.4 thousand (an increase by 6.5% compared to the previous year and 10.7% compared to 2011) with persons hired in ICT services constituting over three quarters. IT services were also the field of activity in which enterprises hired the biggest number of persons of all employed in ICT services (66.1%). The value of net revenues from sales in the ICT sector increased by 8.8% in comparison with 2011 and amounted to over PLN 132 billion in 2014 (a slight decrease was only noted in 2013 compared to 2012 – 0,7%). Services, in particular telecommunications, had the biggest contribution in generating revenues of the ICT sector. In 2014 ICT manufacturing enterprises earned almost two thirds of revenue from export sales while ICT service enterprises – only 15.4%. In 2011-2013 these revenues were decreasing systematically in ICT manufacturing enterprises but in 2014 they recorded an annual growth of 8.3%. In service enterprises revenues from exports were increasing continually. The biggest share in these revenues had enterprises providing IT services (in

2014 – 63.5%). An increase of expenditures on R&D in the ICT sector was noted in the years 2011-2014 (by PLN 553 million). Enterprises offering ICT services incurred about 90% of expenditures on R&D in each surveyed year.

3 Information society

The Government of the Republic of Poland, having in mind the good of Poland and its inhabitants, is striving to ensure rapid and sustainable economic growth and social development that will improve the living conditions of our citizens. One of the key stimulants of economic growth is citizens' ability to acquire, accumulate and use information as a result of the dynamic development of Information and Communication Technologies (ICT). The importance of this factor for economic growth is confirmed by various research projects, which conclude that ICT account for approximately a quarter of the GDP growth and 40% of the productivity growth in the European Union. The rapidly increasing significance of information and electronic services, and thus, of the application of ICT, in the economy, public administration (central and local government) and in the everyday life of citizens has triggered a new transformation trend – a transformation towards the "information society". The term "information society", as adopted for the purposes of this paper, is defined as a society for which the processing of information with the use of ICT solutions creates significant economic, social and cultural value. This Strategy is sectoral and, as such, defines the vision and mission for the development of the information society in Poland until 2013. Within each of its three areas – Human, Economy and State – it maps out strategic directions and determines the objectives that should be accomplished in order to achieve the desired development status for the information society in Poland in 2013. The creation of the Strategy was preceded by a series of extensive consultations with experts who represented organizations and institutions that are most competent to express views on the issue of information society development.

It is possible to distinguish five definitions of an information society, each of which presents criteria for identifying the new. These are: technological, economic, occupational, spatial and cultural. These need not be mutually exclusive, though theorists emphasise one or other factors in presenting their particular scenarios. However, what these definitions share is the conviction that quantitative changes in information are bringing into being a qualitatively new sort of social system, the information society. In this way each definition reasons in much the same way: there is more information nowadays, therefore we have an information society. As we shall see, there are serious difficulties with this ex post facto reasoning that

argues a cause from a conclusion (Webster 2006). Research on the links between the diffusion of Information and Communication Technologies (ICTs) and social and economic development has been undertaken for decades. Evidence of links between social and digital engagement, particularly with respect to the Internet, has been the focus of many studies conducted by academic as well as government institutions. These studies have shown consistently that individuals who have access to ICTs, from the telephone to the Internet, tend to have more schooling, higher incomes, and higher status occupations than do those who do not have access. To analyze this study in the next part of this paper we present statistical analysis based on data from the Central Statistical Office on information society in Poland.

4 The survey results and application in R

In this part of the paper we present multivariate statistical analysis of information society in Poland with the use of R software. We use the data from the report published by the Central Statistical Office entitled *Information Society in Poland*. We present analysis with the use of data on the number of enterprises and employees in the ICT sector in 2011-2014 (table 1). First, we present correspondence analysis for number of enterprises in the ICT sector in 2011-2014. The analysis is based on data from the Central Statistical Office on the number of enterprises in the ICT sector in 2011-2014.

Specification	2011	2012	2013	2014
ICT production	245	239	225	235
ICT wholesale	190	207	230	235
Telecommunications	219	231	258	289
IT services	1070	1181	1305	1387

Table 1. Number of enterprises in the ICT sector. Source: Central Statistical Office.

We present results of correspondence analysis for data presented in table 1.

```
Principal inertias (eigenvalues):
           1        2       3
Value      0.001629 9.3e-05 7e-06
Percentage 94.22%   5.38%   0.4%
```

The percentage being explained by first dimension is 94.22%, and for the second dimension is 5.38%. These two dimension explain 99.6% of the total inertia. Total inertia is

0.0017 showing that there is a very weal association between two variables: year and specification. Row and column masses, chi-square distance and inertia for each of the category of row and column are presented below.

Rows:

	ICT Production	ICT wholesale	Telecommunications	IT services
Mass	0.121869	0.111283	0.128712	0.638136
ChiDist	0.107785	0.015622	0.030357	0.016220
Inertia	0.001416	0.000027	0.000119	0.000168
Dim. 1	-2.669651	0.091573	0.490427	0.394954
Dim. 2	0.207310	-1.431880	2.366810	-0.267273

Columns:

	2011	2012	2013	2014
Mass	0.222566	0.239866	0.260522	0.277046
ChiDist	0.062238	0.023124	0.032682	0.040783
Inertia	0.000862	0.000128	0.000278	0.000461
Dim. 1	-1.532660	-0.524810	0.765233	0.966059
Dim. 2	0.664981	-0.877842	-1.068300	1.230398

We can display the result of a correspondence analysis in the form of perception map (fig. 1).

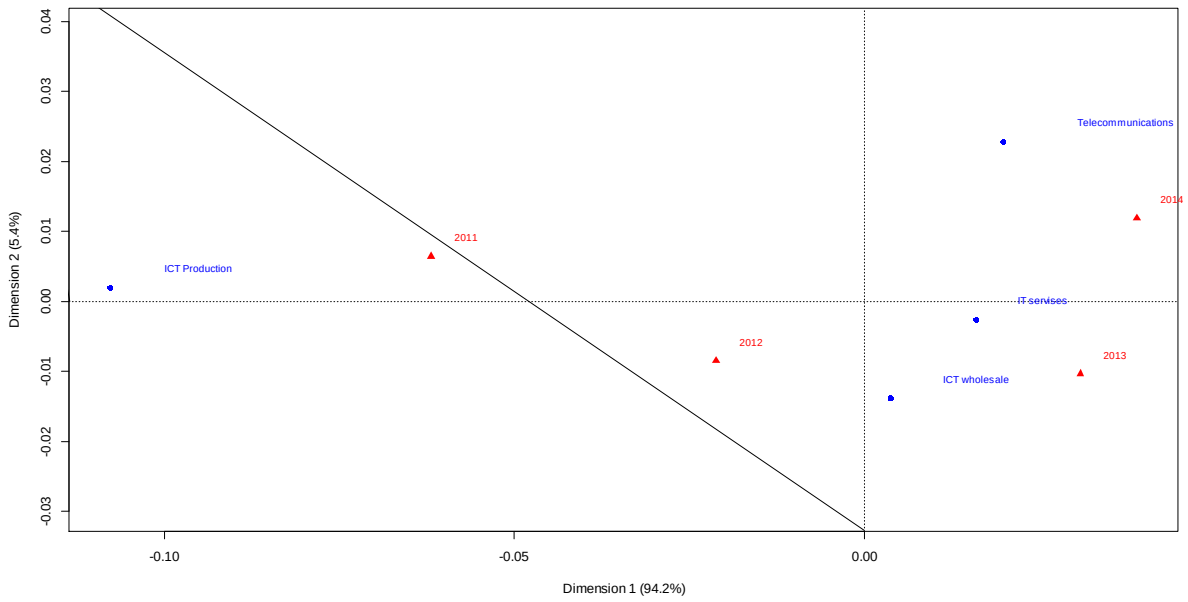

Fig. 1. Two-dimensional perception map for correspondence analysis for the number of enterprises in the ICT sector.

Looking at the graph above (figure 1) we see, the ICT production is situated very close to the year 2011. ICT wholesales and IT services are related with years 2012 and 2013. Finally, telecommunications is situated very near to the year 2014. This graphical presentation may suggest that there is a trend moving ICT production from 2011 to the telecommunications services in 2014.

To present graphically a tree diagram for categories of rows for the number of enterprises in the ICT sector (ITC Production, ICT wholesale, Telecommunications, and IT services) we will apply agglomerative hierarchical clustering. Following is a dendrogram of the results of running these data through the Ward clustering algorithm.

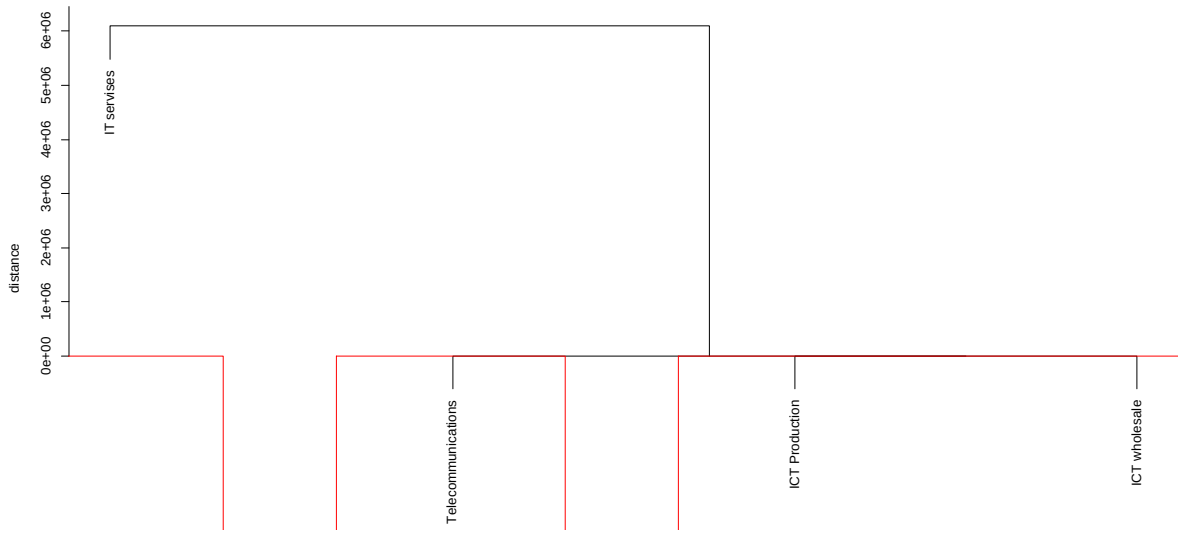

Fig. 2. Dendrogram for the number of enterprises in the ICT sector using Ward method.

We can see, that there are three clusters of ICT specifications separated. One cluster contains two categories: ICT Production and ICT wholesale, second cluster contains only one category - Telecommunications, and last cluster containing also one category is IT services. The analysis is based on data from the Central Statistical Office on the number of employees in the ICT sector in 2011-2014. The correspondence analysis was also conducted for the number of employees in the ICT sector (table 2).

Specification	2011	2012	2013	2014
ICT production	44930	41150	36892	39337
ICT wholesale	10363	10598	11372	11496
Telecommunications	46516	43890	42634	41786
IT services	75539	85178	93422	103739

Table 2. Number of employees in the ICT sector. Source: Central Statistical Office.

The percentage being explained by first dimension is 97.72%, and for the second dimension is 2.27%. These two dimension explain 99.99% of the total inertia. Total inertia is 0.0067 showing that there is a very weal association between two variables: year and specification. Row and column masses, chi-square distance and inertia for each of the category of row and column are presented below.

```
 Principal inertias (eigenvalues):
           1        2        3
Value      0.006547 0.000152 1e-06
Percentage 97.72%   2.27%    0.01%
```

We can display the result of simple correspondence analysis in the form of perception map.

```
 Rows:
        ICT Production ICT wholesale Telecommunications IT servises
Mass          0.219680      0.059321           0.236622    0.484377
ChiDist       0.099928      0.023115           0.076194    0.080015
Inertia       0.002194      0.000032           0.001374    0.003101
Dim. 1       -1.216980      0.110858          -0.921128    0.988340
Dim. 2        1.378733     -1.703701          -1.282066    0.209650

 Columns:
             2011      2012      2013     2014
Mass     0.240035  0.244729  0.249471 0.265765
ChiDist  0.122501  0.028559  0.054708 0.089983
Inertia  0.003602  0.000200  0.000747 0.002152
Dim. 1  -1.513489 -0.351812  0.631948 1.097723
Dim. 2   0.224847  0.119762 -1.578475 1.168343
```

We can display the result of a correspondence analysis in the form of perception map (fig. 3).

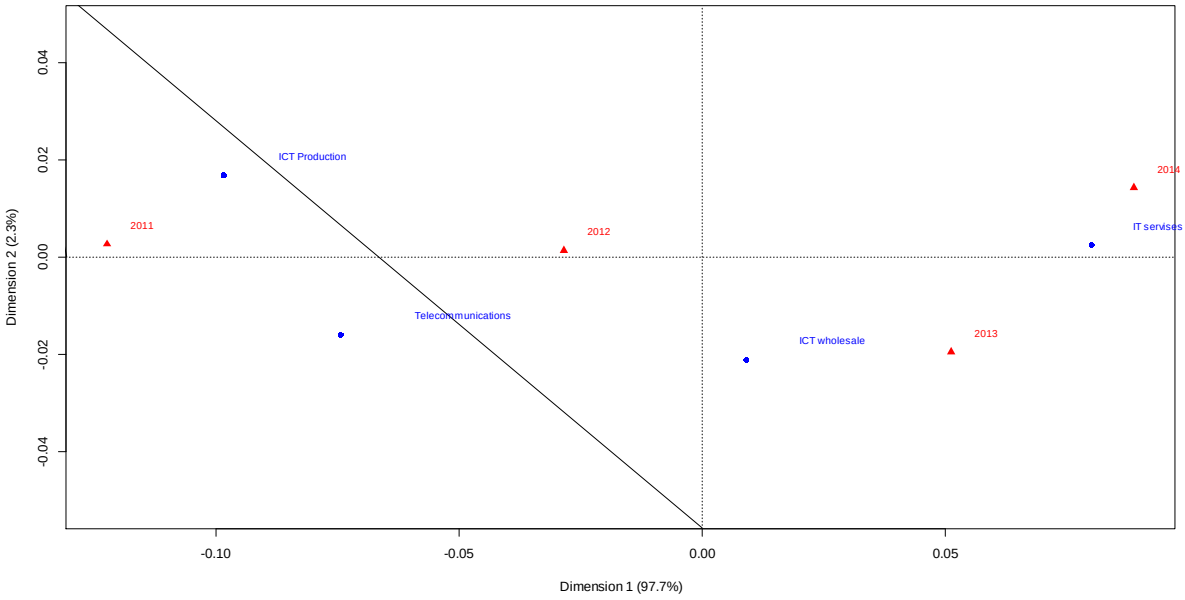

Fig. 3. Two-dimensional perception map for correspondence analysis for the number of employees in the ICT sector.

Looking at two dimensional perception map (fig. 3) we can see that the points are situated in a different way to the analysis conducted for enterprises. The ICT production is situated very close to the year 2011. Telecommunications are related with the year 2012. ICT wholesales is situated very near to the category 2013, and IT services are plotted near the year 2013.

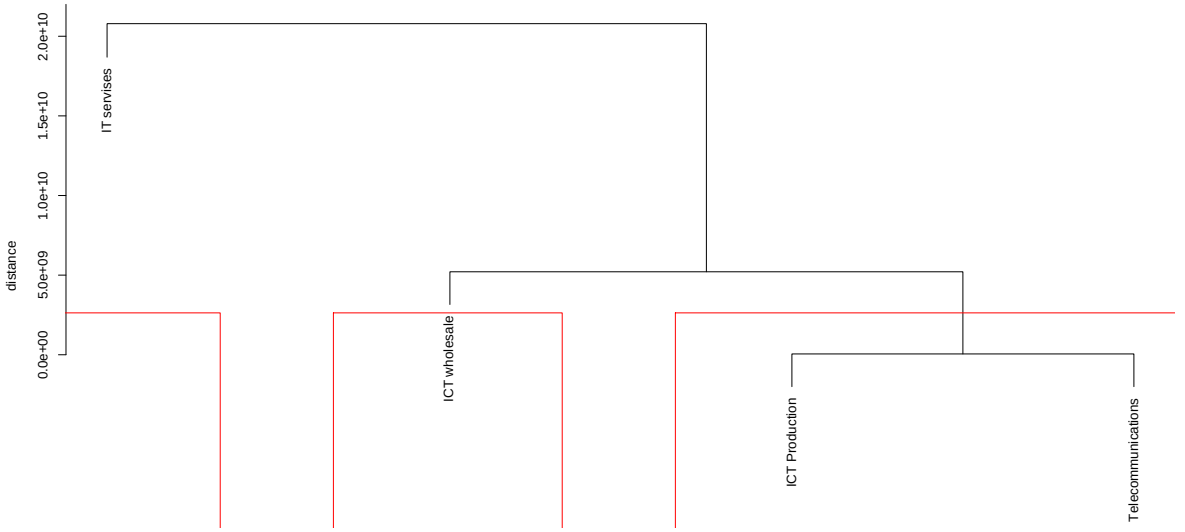

Fig. 4. Dendrogram for the number of employees in the ICT sector using Ward method.

Looking at dendrogram for the number of employees in the ICT sector (figure 4) we can see that there are three clusters separated. First cluster contains two categories: ICT Production and Telecommunications, second cluster contains one category – ICT wholesale, and last cluster contains one category – IT services. Using multivariate methods, we can see, which categories of variables analyzed belong to the cluster of objects that are similar to each other.

Conclusions

Statistical methods applied in the empirical part of this paper allow to see the deeper structure of the number of enterprises in the ICT sector, as well as the number of employees in the ICT sector in 2011-2014. The analysis is based on data from the Central Statistical Office. Statistical methods of the analysis of categorical data were applied, such as: correspondence analysis and hierarchical clustering analysis using Ward method. As a result, a perception map and dendrogram was obtained for a number of enterprises, as well as employees in the ICT sector in 2011-2014. Such methods allow to see which of the categories that were analyzed are similar to each other in the following years. The analysis allows to conclude that busing correspondence analysis for enterprises, we can distinguish three clusters: ICT production, telecommunication, and IT services and ICT wholesale in the same group. For an employee group, there are also three clusters: ICT production, telecommunication, and ICT wholesale with IT services in one cluster. Similar results were obtained in hierarchical analysis. Three clusters were separated: IT services, telecommunication, and the last cluster containing ICT production and ICT wholesale for enterprises data. For an employee's data there are also three clusters, however with different categories: IT services, ICT production, and in the last cluster ICT production with telecommunication. The analysis conducted in the paper shows different business areas that are similar in the years 2011-2014. It also shows the trend in time perspective, which area are more developed and on top in the following years and how time changes the business moves from one area to another.

References

Adams R., & Gill, S. (2007). Augmented cognition, universal access and social intelligence in the information society. In: Schmorrow, Dylan D. & Reeves, Leah M. (eds.), *Foundations of Augmented Cognition*. *Lecture Notes in Computer Science 4565*, Springer, Berlin, 231-240.

Britz, J. (2008). Making the global information society good: A social justice perspective on the ethical dimensions of the global information society. *Journal of the American Society for Information Science and Technology* 59(7), 1171-1183.

Kondratiuk-Nierodzińska, M. (2016). New Knowledge Generation Capabilities and Economic Performance of Polish Regions. Equilibrium. *Quarterly Journal of Economics and Economic Policy*, 11 (3), 451-471.

Piech, K. (ed.) (2007). *Knowledge and Innovation Process in Central and East European economies.* Warsaw: The knowledge Innovation Institute.

Pohle, J. (2015). UNESCO and INFOethics: Seeking global ethical values in the Information Society. *Telematics and Informatics, 32*(2), 381-390.

Sourbati, M. (2011). The digital switchover as an information society initiative: The role of public policy in promoting access to digital ICTs. *Telematics and Informatics, 28*(4), 295-302.

Stephanidis, C. (1998). Toward an information society for all: An international research and development agenda. *International Journal of Human-Computer Interaction , 10*(2), 107-134.

Webster, F. (2006). *Theories of the information society*. Third edition. Routledge / Taylor & Francis c/o Bookpoint. Retrieved February 10, 2017, from http://stat.gov.pl/en/topics/science-and-technology/information-society/information-society-in-poland-results-of-statistical-surveys-in-the-years-20102014,1,1.html