Cluster Analysis of Digital Performance in Educational Techniques in Conditions of EU

Huculova, Eva and Solcova, Lucia

Department of Banking and Investment, Technical University of Košice, Slovakia

Abstract

Global technological trends affect broad spectrum of areas in our life, and through the implementation of particular tools, they are affecting the development of educational levels in particular countries and also educational process itself. This study is aimed at the comparison and revealing of the effect of digitalization and e-skills on the level of education in 20 selected EU member countries by using Factor Analysis and Cluster Analysis. As a result, we consider four clusters of countries with similar characteristics in terms of education, digital literacy and public funding and expenditure on development of ICT and education. Population e-skills demonstrated almost identical levels. The implementation of technological trends into the educational process does not depend only on its positive effects on the educational level, but also on the real opportunity to use these modern tools within the educational process. This is determined by a number of socio-economic, political and cultural aspects. Their detailed examination requires the access to more structured data. Those conditions also create a platform for a subsequent research.

Keywords: DESI index, cluster analysis, technological trends, educational process.

1. Introduction

In the course of several decades, countries around the globe were affected by several system changes, from the change in the policy direction of particular states, to the transformation of economic systems onto market economies, European Union creation, etc. The consequence of these processes is the extension of the economy openness towards global and civilization trends, including technological progress and innovations in many countries. Some countries fall behind in this process; therefore, one of the key preconditions of technological progress is the preparedness of broad range of population classes for using modern information and communication technologies (ICT) or spreading of digital literacy. Although, there is no generally accepted definition of digital literacy (Chetty et al., 2017), according to American Library Association's digital-literacy task force, the digital literacy can be defined as „The ability to use information and communication technologies to find, evaluate, create, and communicate information, requiring both cognitive and technical skills." Digital literacy is the deciding factor of digital transformation. It is the ability to understand the information and use them in various formats presented by means of ICT. It also enables the participation in social networks for creation and sharing of knowledge and supporting wide range of professional computing skills (UNESCO, 2011). Global technological trends affect wide spectrum of areas in our life, and through the implementation of particular tools, they are influencing in some way the development of educational levels in particular countries and also the educational process itself. Although digital technologies have the tendency to improve the teaching process and learning by various methods, they cannot achieve this improvement, as other factors are influencing them. On the other hand, undoubtable advantage is the fact that the digital education can ease and provide education to those, who do not have physical access to it (Pagani et al., 2016; Devaux et al., 2017). An important tool for education are ICT tools which helps students and teachers in the cooperation, communication and by solution of wide spectrum of problems (Afshari et al., 2009). Implementation of technological trends into the educational process is not dependent only by its positive effects on the educational levels, but it depends also on real opportunity to use these modern tools within the educational process. It seems that different level of access to ICT and low level of digital literacy can be a significant future factor of social differences deepening and it can lower chances for quality education (Velšic, 2005; Montoya, 2017). In our paper, we are trying to analyze not only the current level of digitalization in particular EU member country, but also other factors affecting the use of modern trends in the education, where belong mainly government expenses of countries on the education, level of use of ICT in the education, as well as the analysis of households without internet connection due to too high costs. This paper is divided into four parts. In the second part of the paper, we are focusing on the new trends in education that have been transformed by the technology deployment. The third

part describes methodology used in our analysis. Based on our findings and used methodology we have collected and described our results, shown in fourth part. In conclusion, we are summarizing our results.

2. Transformation trends in education

The transformation of educational system is one of the areas of interest and discussions around the globe. European Political Strategy Centre (EC, 2017) points out in its publication 10 main transformation trends in education, displayed on the Figure 1. To new trends in education belong gamification, use of virtual laboratories, role change of a teacher onto mentor and couch, use of interactive aid tools, as well as digitally oriented education using online libraries, interactive tables, webinars.
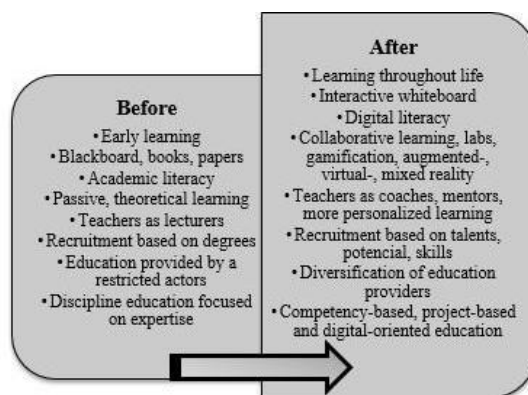

Figure 1. New transformation trends in education. Source: own processing according to EC (2017).

An additional aspect by the application of these transformation trends is the fact that despite the using of technological trends, online available information and materials, it is necessary to establish some culture cooperating on sharing the knowledge and information, and becoming the part of, so called, participative educational culture (Brooks & Gibson, 2012).

3. Methodology

The aim of this paper is the comparison and determination of the digitalization impact and e-skills on the educational level in selected EU countries by means of graphical representation of data through the using geographic information systems (QGIS) and multivariate statistical methods, specifically by factor and cluster analysis. Factor analysis is multidimensional statistical method explaining mutual linear dependence of observed variables by the existence of lower number of unobservable factors called as common

factors (Škaloudová, 2010). The essence of the cluster analysis is formation of clusters, mainly on the basis of similarity, or dissimilarity of clusters, while objects in the cluster are as similar as possible, and clusters are as dissimilar as possible (Trebuňa & Halčinová, 2011; Sebera, 2012). The most frequently used metrics for the calculation of the distance within the Cluster analysis is the Euclidean distance given by the formula:

$$d_{ij} = \sqrt{\sum_{k=1}^{n} \left(X_{ik} - X_{jk} \right)^2}$$

, where X_{ik} and X_{jk} are vectors of similar number of items. *(1)*

In this paper, we apply the cluster analysis by means of Ward´s method. It is necessary to calculate the correlation matrix among particular countries from normalized quantities. Subsequently, we use the statistics Kaiser, Meyer, Olkin (KMO), of which value is higher than 0.5, which proves the correlation in the data and suitability of the use of factor analysis. In this paper, the cluster analysis is created for the year 2015 by the comparison of 20 selected EU countries, while necessary data was obtained from Eurostat databases, United Nations Development Program, EU Data Portal and European Commission. Considered are following variables of the model (Table 1):

Table 1. Description of variables

x1	**Education Index (EI)** – the indicator is calculated from the average index of school education and index of expected years of school attendance.
x2	**Digital Economy and Society Index (DESI) Connectivity** – this indicator is calculated as the weighted average of the four sub-dimensions: Fixed Broadband, Mobile Broadband, Speed, Affordability.
y1	**Percentage of young people using the PC** - Percentage of individuals who used a computer, by 16 to 24 years old.
y2	**Percentage of adults using the PC** - Percentage of individuals who used a computer, by 25 to 64 years old.
y3	**Percentage of young people using the internet** - Percentage of individuals who used the internet, by 16 to 24 years old.
y4	**Percentage of adults using the internet -** Percentage of individuals who used the internet, by 16 to 64 years old.
c1	**Population with secondary education** - The share of population with the highest education by the 25-64 years old, in percentage.
c2	**Population with tertiary education** - The share of population with the highest education by the 25-64 years old, in percentage.
v1	**Expenditures on education** - Government expenditure on education (current, capital, and transfers) as a percentage of GDP.
v2	**ICT Government budget allocations for R&D** – Indicator for distinguishing, which part of budget resources for the research and development sponsors expenses for the area of the ICT development, in Millions of current euros in PPS.

Source: own processing.

4. Results and discussion

Based on obtained data from European Commission and by carrying out analysis through the statistical system QGIS, we can point out to the comparison of the digital performance of 28 EU member countries, where the main examined indicator is a composite index DESI composed of five principal policy areas: Connectivity (25 %), Human Capital (25 %), Use of Internet (15 %), Integration of Digital Technology (20 %) and Digital Public Services (15 %). DESI index uses more than 30 key indicators and it is published every year (EC, 2017). Graphical representation of digital performance of 28 EU member countries is presented by Figure 2, which divides countries by their achieved performance in 2017 in 4 intervals, where the average of EU represents level of DESI index 0.526. The countries with index above the average of EU belong overall 15 analyzed countries (DK, FI, SE, NL, LU, BE, UK, IE, EE, AT, DE, MT, LT, ES, PT), whose composed index DESI is in interval ⟨0.526, 0.707⟩, on the other hand countries with index under average of EU, are 13 (FR, SI, CZ, LV, SK, HU, CY, PL, HR, IT, EL, BG, RO) and their index is in interval ⟨0.332, 0.526). Countries like Denmark, Finland, Sweden and the Netherland lead and the other way around countries like Romania, Bulgaria and Greece are lagging behind in comparison with other analyzed EU countries. Looking at the comparison with last year´s DESI index, countries, which have improved their digital performance this year for more than 8 % are Slovakia (9.9 %), Slovenia (8.9 %), Italy (8.8 %) and Cyprus (8.1 %).
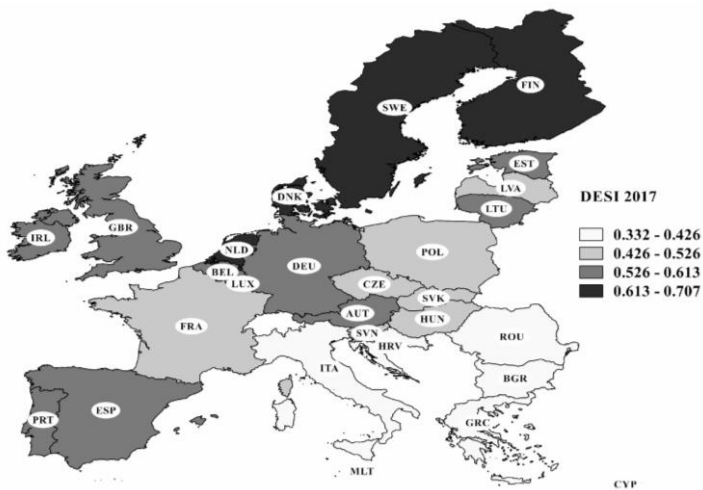

Figure 2. Graphical representation of DESI (Digital Economy and Society Index) of 28 EU member countries, ranking 2017. Source: own processing according to EC (2017).

Secondly, we use multidimensional statistical methods. According to our results, the MSA (measure of sampling adequacy) index of the statistics KMO is higher than 0.5, which means that in this case, it is suitable to use the factor analysis. By the calculation of own

values of correlation matrix (3.98; 1.99; 1.41; 0.77; 0.67; 0.47; 0.44; 0.18; 0.051; 0.018) we have found out that three own numbers are higher than 1; therefore, as suitable number of factors are chosen three factors. From the results of this factor analysis it is clear that the first factor explains almost 40 % of total variability; the second factor explains almost 20 % of total variability and the third factor approximately 14% of total variability. The percentage of remaining factors on the variability is relatively low, so we can state that for the explanation of original variables, it is necessary to use three factors, by which we are able to explain almost 74 % of total variability. Continuing with the factor analysis directly with the rotation VARIMAX, by means of which we can estimate positive or negative correlation of the indicator with one factor and subsequently, we can calculate the factor score. The results of this analysis are three factors and their load called as factor saturation, introduced in the Table 2.

Table 2. Factor load after rotation VARIMAX.

variables	factor 1	factor 2	factor 3
X_1	0.15	0.18	**0.68**
X_2	**0.58**	0.36	0.43
Y_1	0.28	**0.74**	-0.07
Y_2	0.55	0.50	0.53
Y_3	-0.14	**0.89**	-0.09
Y_4	0.06	**0.78**	0.35
C_1	**0.95**	-0.02	0.14
C_2	**0.94**	-0.05	0.06
V_1	**0.78**	0.16	-0.03
V_2	-0.05	-0.19	**0.85**

Source: author´s work.

The Table 2 indicates that after the variation VARIMAX, it is possible to assign each variable (exception of y_2) as positively or negatively correlating with one of three factors. The variable y_2 cannot be explained by any factor. It is slightly positively correlated by all three factors. On the basis of results of factor load, we can interpret the first factor 1 as the factor of the educational level in the connection with government financing (variables of the percentage of the population with secondary and tertiary education, and expenditures on education as the percentage of GDP). Second factor can be interpreted more clearly as digital skills (variables of the percentage of the population using the PC and internet) and the third factor 3 as the level of financing of ICT (variable budget resources for the research and development of ICT). In the last step, we have created the cluster analysis, and on the basis of graphical output (Figure 3), as optimal number of clusters we consider 4 clusters. This statement is also supported by Ratkowsky index used for the determination of optimal number of clusters and its value is 0.4015, by which it is recommended to consider 4 clusters, which is displayed also on the Figure 3.
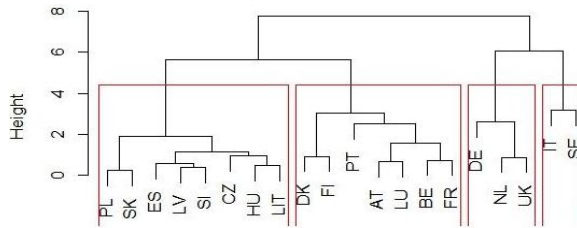
Figure 3. Cluster analysis of EU countries in the form of Dendrogram. Source: own processing.

According to results of the cluster analysis (Figure 3), we have considered 4 clusters or groups of examined countries. The first and the smallest group within the cluster analysis with similar characteristics contain Italy and Sweden. In this cluster, mentioned countries are comparable in the indicator Education index (x_1) with average value 0.830, as well as in the indicator DESI Connectivity (x_2) with average value 0.158. From all results, countries in this group are typical by high percentage of expenses on education (4 – 6.3 % of GDP). Countries of this cluster differ by indicators c_1 and c_2 (with average values 15.57 and 26.60), which represent different level of the population in the given country with secondary and tertiary education, but also the indicator of budget expenses for the development of ICT with average value v_2=216. Second cluster includes countries United Kingdom, Netherlands and Germany. In this group, similar values can be found in the indicator x_1 and x_2 with average values 0.851 and 0.13. Indicator c_1 achieves values 3.39 – 4.68 %. Significant differences in averages of the second group of countries can be found by the indicator v_2 with average value 101.84, as well as by c_1, c_2 with average values 5.375 and 12.837. In the case of the third cluster, which is formed by countries France, Belgium, Luxemburg, Austria, Portugal, Finland and Denmark we can see average values by indicators x_1=0.901 and x_2=0.173. The indicator v_1 achieves the average value 4.34%, which is comparable with the second analyzed group of countries. Also, in this cluster, the most significant differences of particular countries are in indicators c_1=13.033, c_2=20.366, as well as by the indicator v_2=861.39. Fourth cluster belongs to Lithuania, Hungary, Czech Republic, Slovenia, Latvia, Spain, Slovakia and Poland with results x_1=0.835 x_2=0.135. Indicators v_1=4.375, v_2=458.44, c_1=16.55, c_2=27,75 indicate the biggest differences in values within the group, with average values.

5. Conclusion

According to our digital performance comparison and its financing we have calculated four clusters of selected European countries with similar characteristics where the highest values were analyzed in the first and third clusters (DK, FI, PT, AT, LU, BE, FR and DE, NL, UK). In all four examined clusters, indicators are representing population e-skills which have almost identical levels. Living in higher educated countries does not necessary mean higher level of digital performance. The implementation of technological trends

into the educational process does not bring only its potential positive effects on the educational level, but also the opportunity to make educational process easier, more interactive and digitally oriented. Despite of the fact that the amount of internet availability has increased slightly in last years, some countries have still low internet affordability. This is determined by socio-economic, political and cultural aspects. Their detailed examination requires the access to more structured data. Wide development of digitalization in educational system can motivate young generations to support and create the digital single market in Europe.

References

Afshari, M., Bakar, K.A., Luan, W.S., Samah, B.A., Fooi, F.S. (2009). Factors Affecting Teachers´use of Information and Communication Technology. *International Journal of Instruction*, *2*(1), 77-104.

American Library Association's digital-literacy task force. (2012). Digital Literacy Definition. Retrieved from http://connect.ala.org/node/181197.

Brooks, Ch., Gibson, S. (2012). Professional Learning in a Digital Age. *Canadian Journal of Learning and Technology, 38*(2), 1-17.

Chetty, K., Liu, Q., Wenwei, L., Joise, J., Gcora, N., Shenglin, B. (2017). Bridging the Digital Divide: Measuring Digital Literacy. G20 Insights, Germany, 1-8.

Devaux, A., Bélanger, J., Grand-Clement, S. and Manville, C. (2017). Education: Digital technology's role in enabling skills development for a connected world. RAND Corporation, PE-238-CI (2017). https://www.rand.org/pubs/perspectives/PE238.html.

European Commission. (2017). 10 Trends Transforming Education as We Know It. Retrieved from https://ec.europa.eu/epsc/publications/other-publications/10-trends-transforming-education-we-know-it_en.

European Commission. (2017). Digital Economy and Society Index (DESI) 2017. Retrieved from http://europa.eu/rapid/press-release_MEMO-17-352_en.pdf.

Montoya, S. (2017). Tracking Literacy in an Increasingly Digital World. UNESCO Institute for Statistics. Retrieved from http://uis.unesco.org/en/blog/tracking-literacy-increasingly-digital-world-0

Pagani, L., Argentin, G., Gui, M., Stanca, L. (2016). The impact of digital skills on educational outcomes: evidence from performance tests. *Educational Studies*, *42*(2), 137-162.

Sebera, M. (2012). Vícerozměrné statistické metody. Shluková analýza. Masarykova univerzita, Brno. Retrieved from http://www.fsps.muni.cz/~sebera/vicerozmerna_statistika/shlukova.html.

Škaloudová, A. (2010). Faktorová analýza. Faktorová skóre. Univerzita Karlova, Prague. Retrieved from http://kps.pedf.cuni.cz/skalouda/fa/faktorova_skore.htm.

Trebuňa, P., Halčinová, J. (2011). Graphic Interpretation Cluster Analysis and select the number of important clusters. The 14th International Scientific Conference Trends and Innovative Approaches in Business Processes, 1-3.

UNESCO, Institute for Information Technologies in Education. (2011). Digital Literacy in Education. Retrieved from http://unesdoc.unesco.org/images/0021/002144/214485e.pdf

Velšic, M. (2005). Digitálna gramotnosť na Slovensku. Správa z výskumu. Inštitút pre verejné otázky, Bratislava.