

COMPARATIVE ANALYSIS OF THE INDUSTRIAL PRODUCTION INSIDE THE EURO ZONE

Gabriela Ižaríková¹ - Miriam Pekarčíková² - Peter Trebuňa² - Jana Kronová³

Abstract: *In the countries of the European Union the industry is developing rapidly. Industry production has a tendency to grow, but it has different measure. EU countries that are based on similar Industrial production index are divided into certain groups. In this article we analyze the industrial production index in the countries of the European Union on the period 2002-2017 by cluster and regression analysis. Both methods give equal results. The development of the values in the industrial production index have a character growth, but in some countries faster. In all countries there are seasonal variations. According to the trend and seasonal hesitation the EU countries are divided into 4 clusters. Countries that belong to the certain clusters have familiar development of the industrial production index. The fastest growth of the industrial production index is in the countries of the second cluster. The industrial production index in the countries that are in the fourth cluster have declining and constant character.*

Keywords: Industrial production index, Cluster analysis, Regression analysis.

1 INTRODUCTION

The industrial production index is calculated according to the methodology in force in the European Union based on the short-term indicators standard. Its calculation is based on a change in the volume of selected products and on a two-stage weighing system. The index is characterized by a change in industrial production in the current year relative to the base year. Indices are calculated for a certain period, for example, indices romr = 100 (the same period last year), or average month 2010 = 100.

This indicator monitors the development of industrial production index only in industrial production. The data source is the annual structural survey of enterprises, supplemented by the cumulative data from the monthly industrial surveys.

The production index is a business cycle indicator which aims to measure changes in value added at factor cost of industry and construction over a given reference period. It does this by measuring changes in the volume of output and activity at close and regular intervals, usually monthly.

In this article cluster analysis and regression analysis are used for classification of EU countries according the industrial production index. We analyzed monthly data from 2002 to 2017 years. We analyzed the development of the IPP of the 28 countries inside the EU, which information we acquired from the www.eurostat.sk.

2 METHODOLOGY AND OBJECTIVES

This section includes descriptions of the specific objectives and methods. There were primary objectives of this study:

- To identify important trends in the evolution of value the industrial production index.

- To identify countries with similar developments of the industrial production index.
- To organize the countries of the European Union on the groups by cluster analysis.
- To determine the trend in the evolution the industrial production index in the groups.
- To determine the prediction on the Slovakia by regression analysis.

Cluster analysis or clustering's the task of grouping a set of objects in such a way that objects in the same group (called cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). It is a main task of exploratory data mining and a common technique for statistical data analysis used in many fields, pattern recognition, image analysis, information retrieval. Clustering is often one of the first steps in data mining analysis. It identifies groups of related records that can be used as a starting point for exploring further relationships. This technique supports the development of population segmentation models, such as demographic based customer segmentation. Additional analyses using standard analytical and other data mining techniques can determine the characteristics of these segments with respect to some desired outcome. (Bijuraj 2013) The time series of the industrial production index are frequently affected by economy changes and others events that are known to have occurred at a particular point of time. Events of this type, whose timing are known, have been termed interventions by Box and Tiao (1975). Interventions can affect a time series in a several ways. They can change the level, either abruptly or after some delay, change trend, or lead to other, more complicated, response pattern. If we analyze the evolution of time series, we are interested not only in the main development trend of the indicators, but also in the course and intensity of any periodic fluctuations, which these time series present. When working with time series, the data must be

adjusted seasonally. The aim of seasonal adjustment is to uncover the underlying dynamics in the development of the investigated phenomena and allow a direct comparison of their development in different seasons within the year. There are many methods of seasonal adjustment. In the construction of the forecasts of seasonal time series, a regression model with artificial (dummy) variables with simultaneously estimated trend and seasonality parameters can be used. The trend component is modeled via suitable regression function, for example line, parabola, and so on. The seasonal component is expressed using seasonality indexes. (Rublíková 2007). Let us assume an multiplicative time series model. Furthermore, we assume that the time series has a linear trend and monthly seasonality.

3 CLUSTER ANALYSIS

The cluster analysis is overall term for group of methods, which purpose is reduction of set of data on several relatively homogeny groups - cluster, while the condition is maximal and simultaneously minimal

similarity of clusters. Methods of cluster analysis are on the basis of clustering classified as follows:

Hierarchical clustering – are based on hierarchical systematization of objects in cluster. Hierarchical methods leads to hierarchical (tree) structure, which is graphically figured as dendrograph (tree diagram). Dendrograph is graphical output of cluster analysis and it is graphical figuration of cluster on the variate cluster degrees.

Non-hierarchical clustering – is about the categorization of objects into the count of cluster, specified in advance.

Similarity of objects is studied by the degree of similarity (correlation coefficient and association coefficient) or the degree of dissimilarity – degree of distance (distance coefficient). Cluster analysis was accomplished in sw. STATISTICA and Wards method was used. Distance(dissimilarity) is described by means of squared Euclidean distance. Result of hierarchical method of cluster analysis is dendrograph typified in the Fig.1, which depicts clusters of EU countries according to their mutual similarity of the industrial production index.

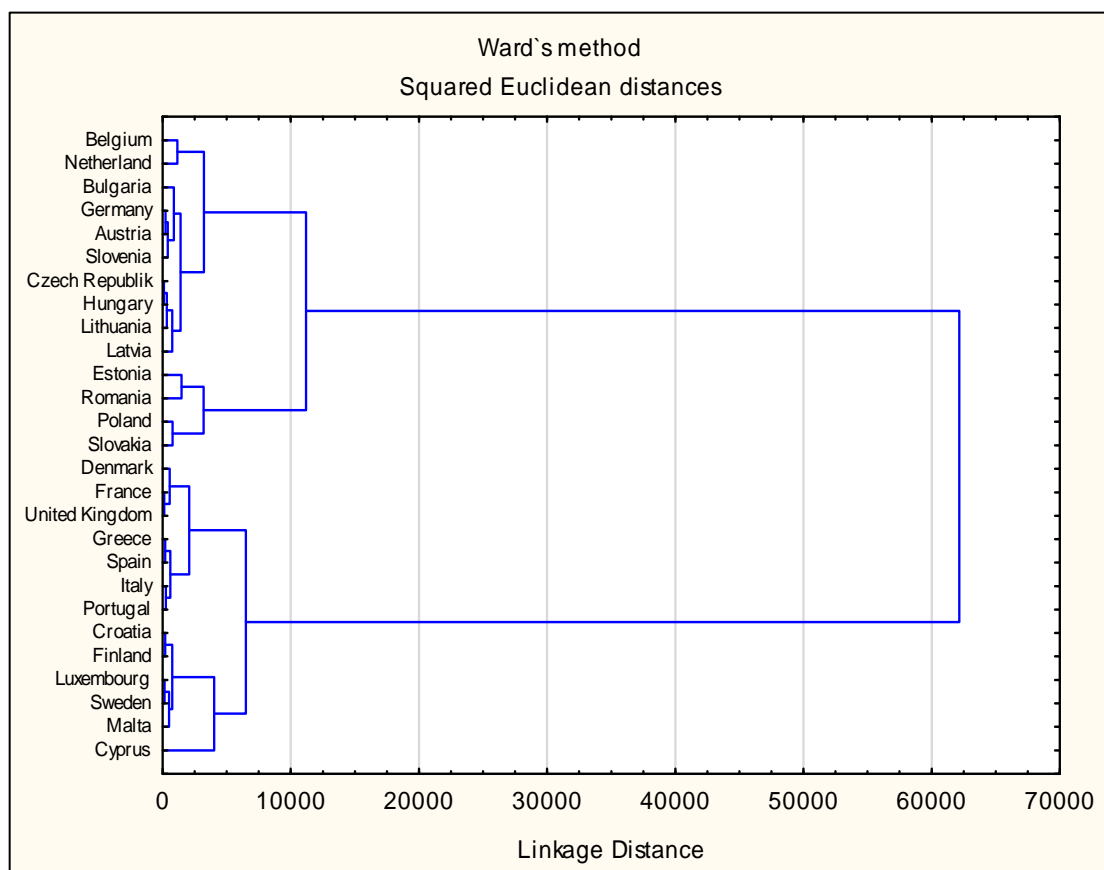


Fig.1 Dendrograph of countries of EU

Table 1 Cluster of countries of EU

1.cluster	2.cluster	3.cluster	4.cluster
Belgium	Estonia	Denmark	Croatia
Netherland	Romania	France	Finland
Bulgaria	Poland	United King.	Luxembourg
Germany	Slovakia	Greece	Sweden
Austria		Spain	Malta
Slovenia		Italy	Cyprus
Czech Republic		Portugal	
Hungary			
Lithuania			
Latvia			

Dendrograph depicts the cut, which represents optimal selection of clusters according to the opinion of authors. We consider following dividing of EU countries into the clusters as optimal, Table 1. In the

following map (Fig. 2) can be seen single EU countries in color discerned according to their belonging to clusters.

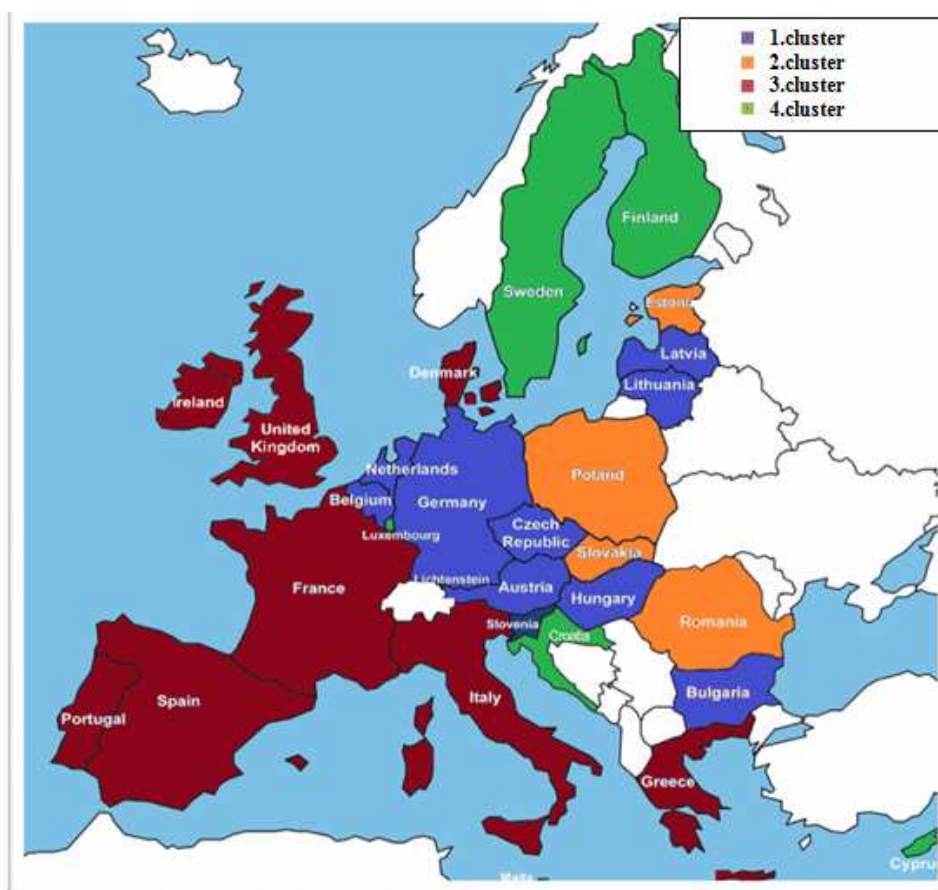


Fig. 2 Maps of clusters

At disposal are data on mutual similarity of single objects (EU countries). In the following Tab. 2 behalf of vast of data and lucidity we only mention data on the smallest distance = the biggest distances of objects, the biggest distance = the smallest similarity of objects (single EU countries).

From analysis results that most similar are according to the industrial production index countries

Hungary and Czech Republic (with the measure of distance 92) and least similar are Slovakia and Greece (with measure of distance 20307). The most distant object from other objects is object – Slovakia.

The following tables mention the countries with the biggest similarity (Tab. 3) and countries with the smallest similarity (Tab. 4) with related values of distance.

Table 2 Comparison of distance

	Biggest similarity smallest distance		Smallest similarity biggest distance	
Belgium	402	<i>Austria</i>	10227	<i>Greece</i>
Bulgaria	545	<i>Slovenia</i>	6584	<i>Cyprus</i>
Czech Republic	92	<i>Hungary</i>	9254	<i>Greece</i>
Denmark	253	<i>France</i>	13219	<i>Slovakia</i>
Germany	219	<i>Austria</i>	6134	<i>Greece</i>
Estonia	1317	<i>Latvia</i>	17052	<i>Cyprus</i>
Greece	202	<i>Spain</i>	20307	<i>Slovakia</i>
Spain	202	<i>Greece</i>	18679	<i>Slovakia</i>
France	128	<i>UK</i>	12417	<i>Slovakia</i>
Croatia	217	<i>Finland</i>	8677	<i>Slovakia</i>
Italy	245	<i>Portugal</i>	15912	<i>Slovakia</i>
Cyprus	1995	<i>Croatia</i>	17451	<i>Slovakia</i>
Latvia	387	<i>Lithuania</i>	9502	<i>Cyprus</i>
Lithuania	237	<i>Hungary</i>	9906	<i>Greece</i>
Luxembourg	161	<i>Sweden</i>	11711	<i>Slovakia</i>
Hungary	92	<i>Czech R.</i>	9539	<i>Greece</i>
Malta	255	<i>Sweden</i>	9249	<i>Slovakia</i>
Netherland	615	<i>Croatia</i>	7213	<i>Slovakia</i>
Austria	126	<i>Czech R.</i>	8173	<i>Greece</i>
Poland	787	<i>Slovakia</i>	18544	<i>Greece</i>
Portugal	245	<i>Italy</i>	16233	<i>Slovakia</i>
Romania	1112	<i>Lithuania</i>	14116	<i>Greece</i>
Slovenia	233	<i>Germany</i>	5219	<i>Slovakia</i>
Slovakia	787	<i>Poland</i>	20307	<i>Greece</i>
Finland	173	<i>Sweden</i>	9224	<i>Slovakia</i>
Sweden	161	<i>Luxembourg</i>	10002	<i>Slovakia</i>
UK	128	<i>France</i>	11612	<i>Slovakia</i>

Table 3 Biggest similarity (smallest distance)

Distance	Countries	
92	Czech Republic	Hungary
127	France	UK
161	Luxembourg	Sweden
202	Greece	Spain
217	Croatia	Finland
219	Germany	Austria
244	Italy	Portugal

Table 4 Smallest similarity (biggest distance)

Distance	Countries	
20307	Slovakia	Greece
18679	Slovakia	Spain
18544	Poland	Greece
17451	Cyprus	Slovakia
17052	Estonia	Cyprus
16233	Portugal	Slovakia
15912	Italy	Slovakia

4 REGRESSION ANALYSIS

Monthly values of the industrial production index are the time series. Time series is a chronological sequence substantially in essence, space and time comparable values of observations recorded at the time. Time series analysis is used to describe dynamic systems and to build predictions for their future development. This one-dimensional model can be written in the form $y_t = f(t, \varepsilon)$, where t is time, y_t is the value of a variable ε is a random component. Standard model is based on the decomposition of the time series into four components:

- Secular trend which operate continuously, long-term and determine the main direction of development. The values of the time series declined or stagnated.
- Periodic which can operate alternately to increase or decrease the values in the time

series. According to their periodicity duration can be divided into:

- Seasonal component describes periodically recurring changes in the time series taking place within one calendar year or season. Seasonal fluctuations are regularly repeated in certain months or quarters as a result of seasonal effects, such as changing seasons.
- Cyclical component describes the periodic changes in more than one year (recurring period in agriculture changing the production cycle, fashion). Length of the cycles is usually non-constant as growth and decline in economic activity is between 5-7 years.

To identify seasonality respectively cyclical component of the time series can be used a variety of methods (periodogram, spectral analysis, autocorrelation function) of which the simplest is to analyse the graph.

Residual components affect random fluctuations over time series. Random irregular components show irregular fluctuations around the trend values that arise as a result of accidental or unforeseen impacts. It is the result of the action of random effects which have not expressed a function of time.

Division of the time series for each component is called decomposition. Presence of all four components in each time series is not required and depends on the nature of the observed effect. Creating an accurate prediction often requires separate summing upon the components ($T_t, S_z, C_t, \varepsilon$). Decomposition of time series into its individual components, may be:

- Additive: $y_t = T_t + S_z + C_t + \varepsilon$
- Multiplicative: $y_t = T_t \cdot S_z \cdot C_t \cdot \varepsilon$

The industrial production index is a business cycle indicator which measures monthly changes in the price-adjusted output of industry. This article takes a look at the industrial production index in the European Union (EU). The IPI values show a certain trend, impact on the development of industry in individual countries but also across the EU.

Since mid-2003 total industrial output was on a relatively steady growth path. The production level reached its highest value in April 2008 and then fell continuously for one year until in April 2009. Afterwards the indicator steadily increased. In the second half of 2011 and 2012, industrial production in the EU-28 was on a slow downward trend. Since early 2013 the index value is again slightly but steadily increasing. The 2009 decline in production was particularly strong for machinery and equipment and for motor vehicles which together account for around one fifth of total industrial production. The economic crisis did not start in all Member States at exactly the same moment. Several countries (Estonia, Greece, Spain, Luxembourg, Malta and Portugal) already recorded rates of negative change in 2008 but a relatively large number of countries still displayed

positive growth rates of industrial production. In 2009 all EU countries experienced a fall in industrial production. In 2010 all EU countries with the exception of Greece, Croatia and Cyprus had returned to positive growth rates. After two years of recovery the EU-28 as a whole in 2012 again displayed a negative industrial development. This downwards trend for the EU-28 continued for the following year. In 2014 the EU returned to a positive growth which continued in 2015, 2016 and 2017.

To compare the development of the industrial production index values, we used a graphical view of the data found. The development trend of the industrial production index of each country divided into four clusters is on the figures 3-6. It is clear from the graphs that the countries assigned to individual clusters show the same or, respectively, very similar trend.

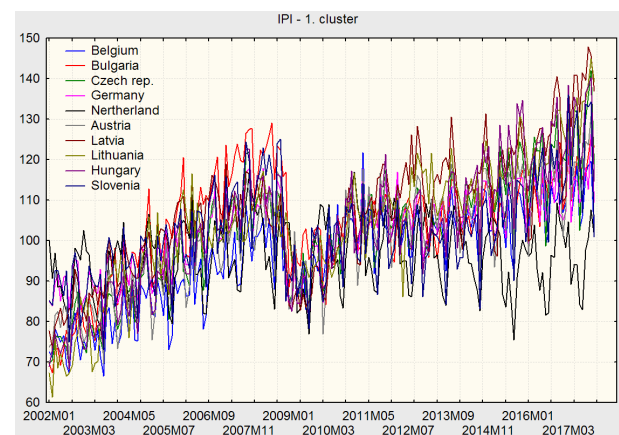


Fig. 3 Index of industrial production – 1.cluster

The IPI values of the first cluster countries from 2002 to 2009 grew, with a decline in 2009, but from March 2003 again the growth rate. In Netherland there is a stronger decline than other countries since 2011. It was the result of the Great Recession of 2008/2009. It contracted again in 2012 due to a relatively low export growth and a drop in domestic spending. The correlation coefficient for the first cluster countries (except Netherland) indicates a significant to high correlation. The locating constant which represents the expected level of the dependent variable at zero value of the independent variable is in the interval (77.39, 93.59) and the regression coefficient which expresses how many units change the dependent variable when changing the independent variable by one unit of measurement is positive in the range of 0.02 to 0.25.

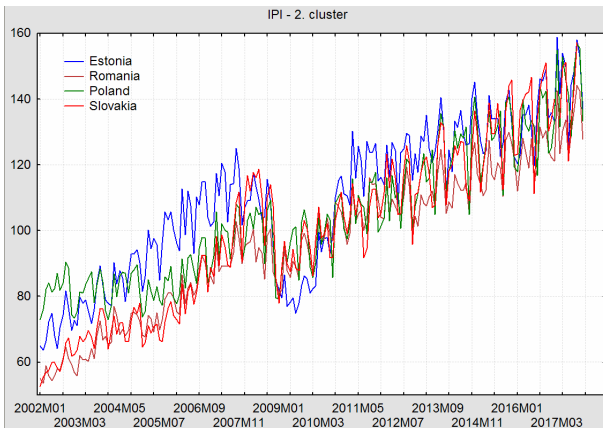


Fig. 4 Index of industrial production – 2.cluster

The IPI countries of the second cluster, including Slovakia, are growing, but in 2009 there is a clear decline caused by the crisis of sales. After 2010, industrial production shows growth, whose dynamics slowed down in 2016. The correlation coefficient for countries of the second cluster shows a high correlation. The locating constant ranges from 59.26 to 75.42 and the regression coefficient ranges from 0.34 to 0.44. This means that monthly IPI may increase by 0.44 but in the first cluster countries by only 0.25.

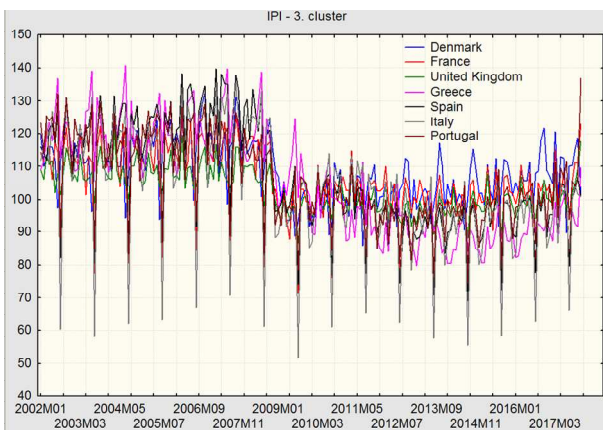


Fig. 5 Index of industrial production – 3.cluster

The third cluster IPI values have a decreasing character the regression coefficient is negative in the range of -0.23 to -0.07 and the locating constant ranges from 109.82 to 127.21. The regression coefficient for Greece points to high tightness (-0.8 indirect dependence), for other countries, it is a slight or significant dependence. The biggest fluctuations are in Italy, the record low values were repeated in August.

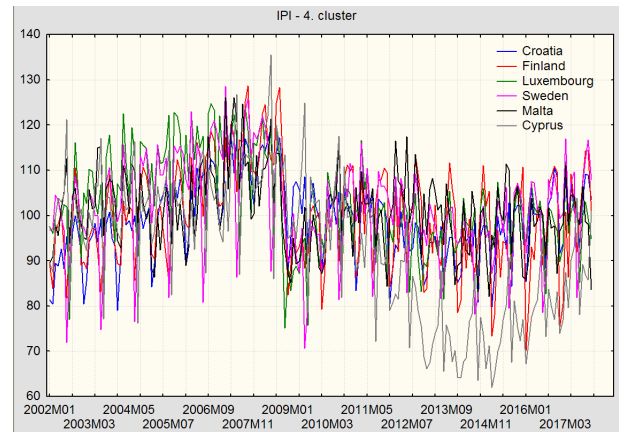


Fig. 6 Index of industrial production – 4.cluster

The fourth cluster IPI values have a slightly declining, rather constant trend, the regression coefficient is in the range -0.18 to 0.00. The locating constant for these countries ranges from 99.45 to 110.07. Dependency for all countries except Cyprus is low. IPI values are very different, oscillated, indicating the instability of industrial production in these countries.

Evolution industrial production index of Slovakia

The data used in this study are sampled on a monthly basis over the period from January 2002 to December 2017, we have time series of monthly values. We assumed model of time series with multiplicative decomposition without a cyclic component and linear trend (see Fig.7). The IPI value graph is supplemented by the 2018 forecast. The prediction took into account the seasonal hesitation and trend of IPI development.

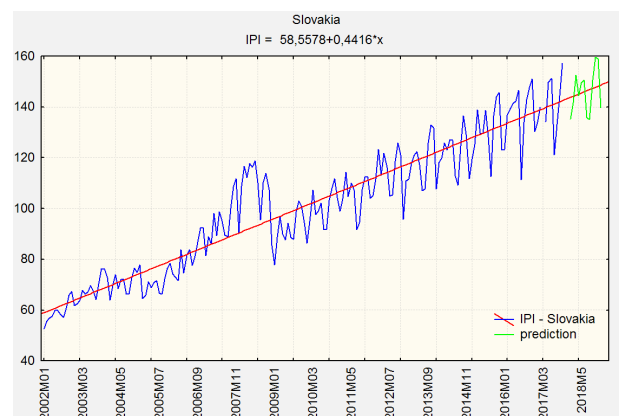


Fig. 7 Index of industrial production – trend and prediction (Slovakia)

The graphical analysis of IPI values is clear of their growing trend of development and season of fluctuations. The evaluation of the trend and suitability of the proposed model is in Table 5.

The regression equation describing the trend has a shape:

$$IPI = 58.55777 + 0.44161n$$

Table 5 Summar regression results

N=192	Regression Summary for Dependent Variable: IPI R=.9348752 R2=.87398850 Adjusted R2=.8733258 F(1,190)=1317.8 p<0.0000 Std.Error of estimate:9.3427					
	b*	Std.Err. of b*	b	Std.Err. of b	t(190)	p-value
Intercept			58.55777	1.353788	43.25474	0.00
n	0.934874	0.025753	0.44161	0.012164	36.30150	0.00

The regression coefficient (0.44161) indicates how many units change the dependent variable (IPI) when changing the independent variable (by one month). The p-value of the significance t-test indicates that both parameters are statistically significant. The determination coefficient (R2 = 0.8733) indicates that the IPI variability is 87.3% explained by this model.

Different statistical methods are used to regulate seasonal hesitation:

- Absolute difference method.
- Relative difference method.
- Calculation seasonal indexes.

We calculated seasonality indexes, which are in the Table 6.

Table 6 Seasonality indexes of IPI (Slovakia)

n	1	2	3	4	5	6	7	8	9	10	11	12
Sz _n	0.94	0.98	1.05	1.00	1.03	1.03	0.93	0.92	1.02	1.08	1.07	0.94

CONCLUSION

Growth of the living standards of the population in the certain countries is affected by development of the economics which is also affected by IPP. This article was focused on application of cluster and regression for analysis of the industrial production index of EU countries. The first part of article is described to the division of countries that are divided into four clusters. Separation into individual clusters is also confirmed by regression analysis methods. Countries in individual clusters show the same trend of development and changed of values. The values of IPI have a growing upward character, but in some countries it is reflected in the aftermath of the crisis and the unfavourable economy.

REFERENCES

- [1] Abonyi, J., Feil, B.(2007) Cluster Analysis for Data Mining and System Identification. Berlin: BV.
- [2] Aldenderfer, M.S.(1984) Cluster analysis. Sage Publications. ISBN 0-8039-2376-7.
- [3] Bacher, J., Poge, A., Wenzig, K. (2010). Cluster analysis - Application-oriented introduction to classification methods, Munchen, Oldenbourg.
- [4] Bijuraj L.V. (2013). Clustering and its Applications, Proceedings of National Conference on New Horizons in IT - NCNHIT 2013
- [5] Everitt, B.S., Landau, S., Leese, M., Stahl, D. (2011). Cluster Analysis, London, Wiley.
- [6] Han, J., Kamber, M. (2006). Data Mining – Concepts and Techniques, San Francisco, MK Publisher.
- [7] Kaufmann, L.(2005). Finding groups in data: an introduction in cluster analysis, Hoboken, Wiley.
- [8] Morrison, S.J. (2009). Statistics for Engineers: an Introduction, Wiley, Chichester.
- [9] Palumbo, M., Lauro, C.N., Greenacre, M.J. (2010). Data Analysis and Classification. Berlin: Springer.
- [10] Rublíková, E.(2007). Analýza časových radov. EKONÓMIA Bratislava .
- [11] Straka, M.; Malindzak, D. (2005) Distribution logistics, TU Kosice, Kosice
- [12] Straka, M. (2013). Logistics of distribution, How effectively to put product into the market, 1st edition, Bratislava, EPOS.
- [13] Trebuňa, P., Halčinová, J.(2012). Experimental modelling of the cluster analysis processes, In: Procedia Engineering, Vol. 48, p. 673–678.
- [14] <http://www.eurostat.sk/>

Acknowledgements: *This article was created by implementation of the grant project VEGA 1/0708/16: “Development of a new research method for simulation, assessment, evaluation and quantification of advanced methods of production.” and project KEGA 030TUKE-4/2017 “Implementation of innovative instruments for increasing the quality of higher education in the 5.2.52 Industrial Engineering field of study”.*

AUTHORS ADDRESSES:

- ¹ Mgr. Gabriela Ižaríková, PhD.
Department of Applied Mathematics and Informatics
Faculty of Mechanical Engineering TUKE
Letná 9, 042 00 Košice
E-mail: gabriela.izarikova@tuke.sk
- ² Ing. Miriam Pekarčíková, PhD.
Institute of management, industrial and digital engineering
Faculty of Mechanical Engineering TUKE
Letná 9, 042 00 Košice
E-mail: miriam.pekarcikova@tuke.sk
- ³ Prof. Ing. Peter Trebuňa, PhD.
Institute of management, industrial and digital engineering
Faculty of Mechanical Engineering TUKE
Letná 9, 042 00 Košice
E-mail: peter.trebuna@tuke.sk
- ⁴ Ing. Jana Kronová, PhD.
E-mail: halcinova.janka@gmail.sk