CLASSIFICATION OF LABOR MARKETS IN THE CZECH REPUBLIC

JAROSLAV URMINSKÝ

ABSTRACT
Geographic segmentation can be admitted in case of assuming the existence of barriers on the labor market. This may result in limitations of the labor market in a certain region with its specific characteristics and possible manifestations of significant inter-regional disparities. The fundamental task is to identify the territorial units facing the negative trends on the labor market. Cluster analysis can serve as one of the instruments for the identification of problems of the regions. Therefore the goal of the paper is to apply selected methods of hierarchical cluster analysis of data in order to identify and group problem areas in the Czech Republic, in the context of the labor market.

KEY WORDS
Labor market, cluster analysis of data, Ward´s method, Czech Republic.

DOI: 10.23762/fso_vol5no1_9

JAROSLAV URMINSKÝ
e-mail: jaroslav.urminsky@vsb.cz

VŠB –Technical University
of Ostrava, Czech Republic

Introduction

Economic theory includes labor among the basic factors of production, which contribute to the production of goods and services. However, work unlike land and capital has specific characteristics. Specific characteristics result from the essence of work, when workers are people with their rights and with psychological and biological characteristics. At the same time employment income is a crucial source of income for most households (Jurečka et al. 2009; Pagés, Stampini 2009). Because of the significant specifics, the labor is the traditional subject of research not only among economists. Among other disciplines that deal with this issue we can include for example, geography, demography, sociology or psychology. In economics, we labor market issue can be found in the most textbooks of macroeconomics and microeconomics (e.g. Samuelson, Nordhaus 1991; Mankiw 1999; Layard et al. 2005; Mach 2001; Holman 2010; Jurečka et al. 2013).

In case, that we do not accept perfect competition in the labor market, which belongs to pure abstraction, there may exist barriers in the labor market. As a consequence of the barriers on the labor market, geographic segmentation of the labor market can appear. Limited spatial mobility of the workforce can be the cause of the geographic segmentation. This relative immobility is attributable in particular to the existence of the transport costs and loss of time in vehicles. In case of migra-

Paper received: 28 May 2016 • Paper revised: 14 December 2016 • Paper accepted: 22 December 2016

tion, migration costs, are considered as important factors. Further, risks caused by unfamiliarity with the environment and psycho-social relation to the original territory (e.g. Hanson, Pratt 1992; Tvrdý et al. 2007). Furthermore, it is reported, that despite currently decreasing geographic immobility, particularly through the development of automobile transport and the development of information and communication technologies, labor market remains in general geographically segmented. As a results labor market is restricted to certain territories or regions. Labor market issues in the regional context are analyzed by number of authors, e.g. Decressin and Fatás (1995), Hančlová et al. (2002), Huber (2006), Tvrdý (2008) and Yonker (2016).

When we accept the premise of geographic segmentation on the labor market, we may try to categorize these markets, respectively classified according to their characteristics. Cluster analysis of data is one of the possible methods to achieve this classification. Meaning of the cluster analysis is to facilitate information retrieval. This is a group of research methods focused on exploratory data analysis, which are included in multivariate statistical methods. The main principle of the cluster analysis is based on decomposition of file into several relatively similar subsets, or clusters. The units within individual clusters should be as similar as possible and the units belonging to different clusters the most dissimilar. Simultaneously, clusters may be different in character. Formed clusters should be preferably compact and mutually disjoint. The cluster analysis is used primarily for classification of objects, but can be used also for reducing the number of variables, where some variables considered are replaced by one variable. In this paper the attention is focused on the classification of objects. The result of the analysis depends on the selected variables, applied measure of distances between objects and clusters and selected calculation algorithm (Řezanková et al. 2007; Kaufman, Rousseeuw 2005; Klímek 2008; Žák 2004 c).

The cluster analysis of data can be treated as widely applicable and appropriate method for assessing the diversity of the regions from the spatial point of view. Application of the cluster analysis provides information value. On the one hand it gives the user a heightened state of knowledge about the territory and also serves as a valuable input in the decision-making process (Kutscherauer et al. 2010).

1. Objective and methods

We used cluster analysis of data for classification of regional labor markets. Cluster analysis can be divided according to type of clustering to hierarchical and non-hierarchical. In the contribution we deal with hierarchical cluster analysis, when using Ward's method.

In case of hierarchical cluster analysis, there are two approaches monothetic and polythetic. We use polythetic approach, which always takes into consideration all the variables simultaneously. In case of this procedure, we create each object in separated cluster at the beginning. In the further steps, there are connected the most similar clusters. The first cluster consisting of two objects based on a matrix of (dis)similarity. Further there are used agglomerative algorithms (distance) for the determination of clusters. The actual procedure of clustering can be illustrated by dendrogram (Řezanková et al. 2007; Romesburg 2004; Meloun, Militký 2002).

The objective of this paper is to apply the selected method (Ward´s) of hierarchical cluster analysis of data, with the intent to identify and segment problem territories in the Czech Republic from the perspec-

tive of the labor market. The territories are defined purposefully by administrative border districts of the Czech Republic outside the capital city of Prague. Thus, these are not natural regions.

As input data we selected indicators characterizing the labor market. The data structure is limited by available of data at the level of districts in the Czech Republic. Here are included only indicators that can be calculated based on the same methodology in all districts of the Czech Republic. Individual districts are treated as objects, where the total number of objects is N = 76. The number of indicators (variables)), p = 8, or p = 9, when the variable no. 9 "Region" not included in the table, indicates belonging of the districts in particular region (Table 1). In the table there are also included the source of data, year, units and direction of action. Direction of action reflects the relation to the labor market. Symbol "–" indicates negative relationship, so higher values are worse. Conversely, "+" symbol expresses the relationship is positive, therefore higher values are better.

Table 1. The structure of the input data

No.	Indicator	Data source	Year	Units	Direction of action
1	Share of unemployed	MoLSA	2013	%	–
2	Number of applicants per 1 vacancy	MoLSA	2013	number	–
3	Number of applicants with higher education per 1 vacancy for undergraduates	MoLSA	2013	number	–
4	Share of long-term unemployed (over 12 moths) in total unemployed	MoLSA	2013	%	–
5	Share of economically active population of working age (15-64 years)	CSO	2011	%	+
6	Share of commuting to work outside the district (another district in the region + other regions + abroad) in total commuting to work in the district	CSO	2011	%	–
7	Number of registered units per 1000 inhabitants	RIS	2013	number	+
8	Total population increase per 1000 mean population	CSO	2013	‰	+

Source: own elaboration based on data from MoLSA 2016 a, b, c, d; Czech Statistical Office 2016 a, b; RIS 2016 a, b

The relevant calculation procedure can be divided into 4 phases in this case:

1.1. Verification of the dependence of variables

To verify the presence of the interdependence of the variables bivariate correlation coefficient (Pearson) was selected. This coefficient takes the values $|0;1|$, with the stipulation that the closer one is dependence higher. The values are heavily dependent, when $X > |0,9|$. In that case all values of variables were $X < |0,9|$, thus no variable was excluded.

Then, there was examined the test of multicollinearity. There was observed Variable inflation factor (VIF). Meloun and Militký (2002) report, that when the VIF>10 the data present strong multicollinearity. For all the variables the values VIF<10, a strong multicollinearity was not revealed.

1.2. Standardization of values

Due to the structure of selected variables with different type of units variables have to be standardized. In case, that we do not make the standardization of data, results of analysis would be most depend-

ent on the variable with the largest range. To the standardize values, the standardization of the z-score is normally used:

$$z_{il} = \frac{x_{il} - \bar{x}_l}{s_l}, \qquad (1.1.)$$

where z_{il} are standardized values, x_{il} – i-th value of the l-th character, $\bar{x}_l$ – the average value of l-th character, s_l – standard deviation of l-th character. A new variable z_l has mean 0 and standard deviation 1 (Řezanková et al. 2007).

1.3. Selecting measure of distance

The literature offers a large number of coefficients intended to measure of the similarity. Choosing the measure of similarity is dependent on the objective of analysis and also on the type of variables. This work concerns the clustering of objects. To the detecting of similarities between objects, we can use the measure of similarity, but more often dissimilarity measure due to their simpler application. Determining the (dis)similarity of objects serves as a decision criterion for the formation clusters of objects (Řezanková et al. 2007; Žák 2004a). Measure of dissimilarity of objects x_i and x_j can generally be written as $D(x_i, x_j)$ possibly D_{ij}.

The selected measure of distance is square Euclidean distance D_{ES}, which is commonly used for application of Ward's method:

$$D_{ES}(x_i, x_j) = \sum_{l=1}^{m} (x_{il} - x_{jl})^2 . \qquad (1.2.)$$

1.4. Selecting clustering method

Distribution into clusters was performed using Ward's method, for which the algorithm between g cluster and unification of clusters h and h' can be written as follows (Řezanková et al. 2007; Žák 2004b):

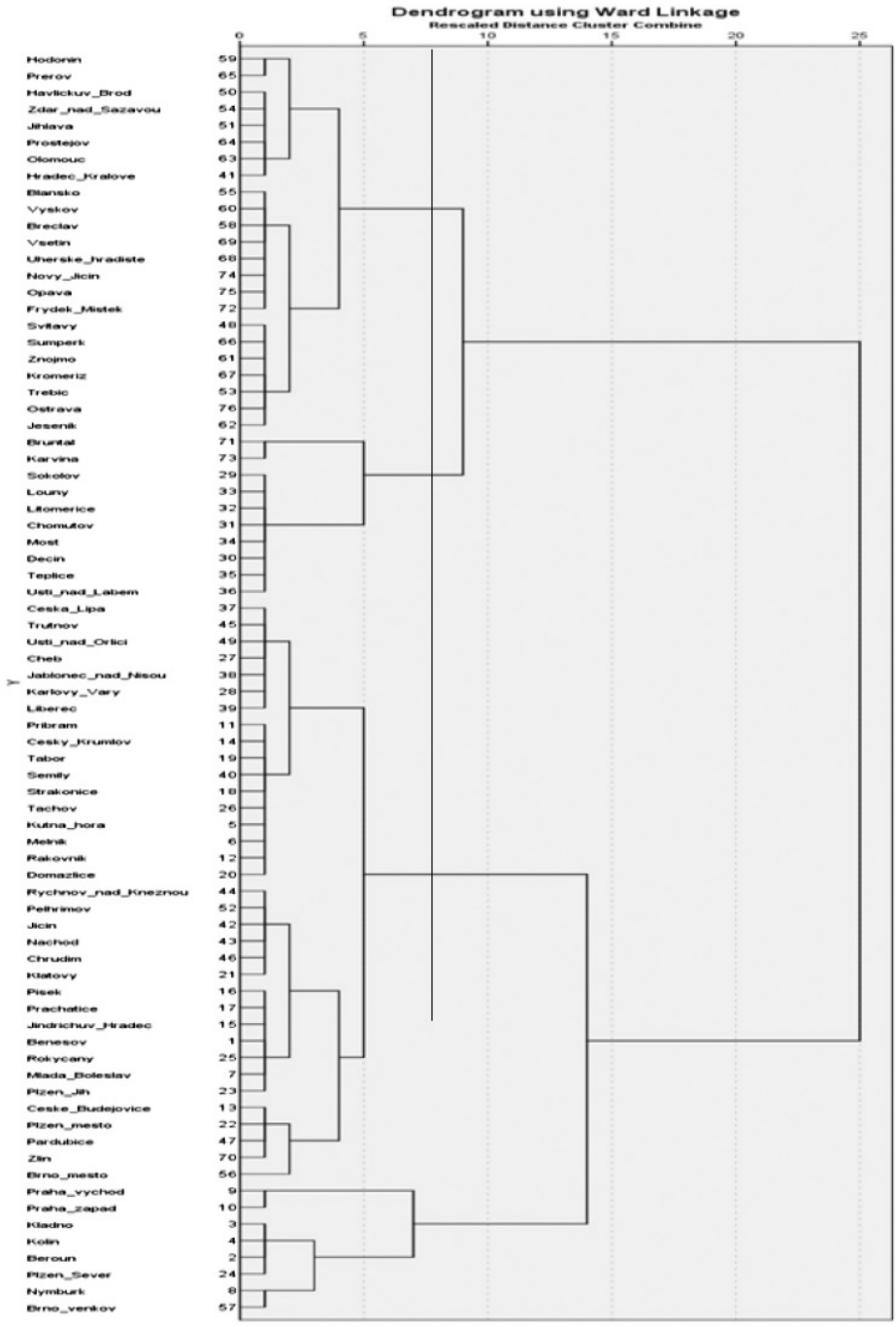
$$D_{g\{h,h'\}} = \frac{(n_h + n_g)D_{gh} + (n_{h'} + n_g)D_{gh'} - n_g D_{hh'}}{n_h + n_{h'} + n_g}. \qquad (1.3.)$$

For the all calculations Microsoft Excel and IBM SPSS Statistics version 22 was used.

2. Classification of the labor markets

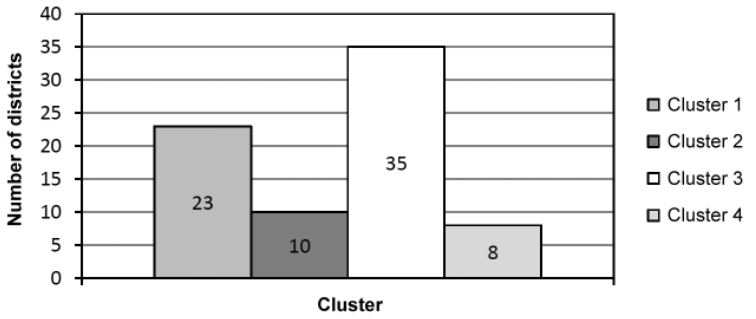
This part is focused on the classification of the labor markets in the Czech Republic. The method used is the Ward's method. Selection of the final number of clusters is made based on heuristics. We elected four final clusters on the basis of resulting dendrogram (Figure 1). Interpretation of the resulting clusters is performed, that the finally clusters we see as separated objects. We proceed to the interpretation of individual clusters with the help of descriptive statistics (Annex 1). In this case there are observed mean values of the variables in the individual clusters. Individual clusters are different in character and in the number of assigned districts (Figure 2 and Annex 2).

Figure 1. Dendrogram – Ward´s method



Source: own elaboration.

Figure 2. Resulting clusters



Source: own elaboration.

Cluster 1 – this is the territory with negative trends on the labor market. We can observe a relatively higher share of unemployed and long-term unemployed. The high number of applicants per vacancy, including the highest number of candidates with higher education per vacancy corresponding to their education is characteristic in this case. The distinctive element is the low level of entrepreneurial activity. The territory is struggling with depopulation and with average share of economically active population. In these territories there is rather low rate of labor commuting outside the district.

Territory – these are mainly districts in the eastern part of the country. All districts in the Olomouc Region. In the Region Vysočina districts Havlíčkův Brod, Žďár nad Sázavou, Jihlava and Třebíč. Kroměříž, Uherské Hradiště and Vsetín in the Zlín Region and districts of Znojmo, Blansko, Vyškov, Břeclav and Hodonín in South Moravian Region. In the Moravian-Silesian Region there belong Opava, Nový Jičín, Ostrava and Karviná. It includes also the districts of Svitavy and Hradec Králové.

Cluster 2 – it is territory with the largest problems from the perspective of the labor market, on the basis of the monitored indicators. A high share of unemployed people and the high share of long-term unemployed, accompanied by low chances of the unemployed to find a job is typical. We see there the lowest level of entrepreneurial activity, significant decrease in population and the low proportion of economically active population. Conversely, university educated workforce here has a relatively high chance of finding a job corresponding educational attainment.

Territory – all districts of Ústí nad Labem Region and the districts of Sokolov, Karviná and Bruntál.

Cluster 3 – the territory can be characterized by relatively low share of unemployed and long-term unemployed. The unemployed have higher chance of finding a job. All is supported by high level of entrepreneurial activity. It is territory with basically stable population and with higher level of economic activity. Share of commuting to the work outside the districts is essentially identical to the clusters 1 and 2, approximately 11%.

Territory – There are all the districts of Liberec and South Bohemian Region. Tachov, Domažlice, Plzeň, Plzeň-Jih, Klatovy and Rokycany in the Plzeň Region and districts of the Karlovy Vary Region, Karlovy Vary and Cheb. In the Hradec Kralove Region we can include there Jičín, Trutnov, Náchod and Rychnov nad Kněžnou and in the Pardubice Region districts of Pardubi-

ce, Chrudim and Ústí nad Orlicí. Rakovník, Příbram, Mělník, Mladá Boleslav, Benešov and Kutná Hora in Central Bohemia Region and districts Pelhřimov, Brno – city and Zlín. In principle there are included especially districts in the western part of the country, with the exception of Brno-city and Zlín.

Cluster 4 – this is the territory with the lowest share of unemployed and long-term unemployed. Job seekers have relatively high chances to find the job, which is not the case of university-educated workforce. Also, there is a high rate of entrepreneurial activity. The territory has significant positive population growth and the highest proportion of economically active population. We observe specific problems with considerable difficulty in finding jobs for university-educated population corresponding to their educational attainment and in the significant share of commuting to work outside the district of residence.

Territory – districts in the hinterland of big cities, Prague-East, Prague-West, Kladno, Kolín, Beroun, Nymburk, Plzeň-sever, Brno-venkov.

Consequently, it is necessary to verify the differences between clusters. To verify the differences between the individual clusters one-way analysis of variance was used. In this case, it was confirmed statistically significant differences between the individual clusters (Annex 3).

3. Typology of regions

We focused on the labor market in the previous part. Based on the dendrogram there were selected four resulting clusters. Based on the comparison of average values of final clusters, we can proceed to determine the typology of regions in terms of the labor market. Interpretation of the resulting clusters is carried out in the way that the final clusters can be seen as separated objects. Depending on the character, one can divide final clusters into:

- stable territory – This territory is the least problematic, characterized by:
  - low share of the unemployed and long-term unemployed,
  - relatively high chances of candidates find a job, including university-educated candidates,
  - high level of entrepreneurial activity,
  - higher level of economic activity of the population,
  - lower proportion of commuting to work outside the district,,
  - slight population decrease.

  The stable region corresponds in the case of Ward's method to the cluster no. 3.
- territory with specific problems – This is the territory with the best results in many aspects. However, there can be identified two specific problems. The territory is characterized by:
  - the lowest share of unemployed and the long-term unemployed,
  - low total number of applicants per one vacancy,
  - high level of entrepreneurial activity,
  - the highest proportion of economically active population,
  - significant population growth.

  Specific problems can be seen in the significant proportion of commuting persons in employment outside the district and in the relatively high number of candidates with university education to 1 vacancy designed for university educated candidates. The territory corresponds in the case of Ward's method to the cluster no. 4.
- vulnerable territory – this is the territory with negative trends on the labor market. The territory is characterized by:
  - higher share of the unemployed, including also the long-term unemployed,
  - higher number of applicants per one vacancy,
  - the highest number of candidates with university education to 1 vacan-

cy intended for university-educated,
- low level of entrepreneurial activity,
- average share of the economically active population,
- relatively low level of labor commuting outside the district,
- slight population decline.

The territory corresponds in the case of Ward's method to the cluster no. 1.

- problem territory - this is the territory with the largest problems from the perspective of the labor market. The territory is characterized by:
  - the highest share of unemployed people and the long-term unemployed,
  - the highest total number of applicants per one vacancy,
  - the lowest number of candidates with higher education on one vacancy,
  - the lowest level of entrepreneurial activity,
  - the lowest share of economically active population,
  - the lowest share of commuting to work outside the district,
  - significant population decline.

The territory corresponds in the case of Ward's method to the cluster no. 2.

Conclusions

For the classification of objects, we used hierarchical cluster analysis of data. The basic objective of cluster analysis is to divide objects into groups so that the objects in those clusters are similar to each other as much as possible and at the same time individual clusters among themselves are most dissimilar. The similarity between objects is essentially a function of their characteristics, although it is a good deal of simplification of reality. Basic characteristics provide information or file of information. These information indicating the character, the essence of observed objects, phenomena or processes may be spatially differentiated. In principle they also implicitly include the characteristics of the population in a defined territory, in this case the districts of the Czech Republic.

The hierachical cluster analysis was carried out on the basis of pre-selected indicators related to the labor market. Ward's method was used. There were identified four clusters distinct in their character. Description of the character and the final clusters can we find in section 2. These clusters form the resulting classification, or the typology of regions. The final clusters were named as stable territory, territory with specific problems, vulnerable territory and problem territory.

Based on the selected indicators Stable territory can be characterized as relatively least problematic. The stable territory consists of 35 districts. There was significant domination of the western part of the Czech Republic. The Territory with specific problems have the best results in many aspects in the context of the labor market. Nevertheless, we identify two specific problems. Firstly, significant proportion of commuting persons in employment outside the districts, in average almost 30%. Secondly, we can see relatively high number of candidates with university education to 1 vacancy intended for university educated candidates. Territory consists of 8 districts. These are districts in the hinterland of big cities. The Vulnerable territory can be characterized by negative trends on the labor market. Territory consists of 23 districts. These are districts mainly in the eastern part of the Czech Republic. The problem territory is territory with the largest problems from the perspective of the labor market. It is characterized by the highest share of unemployment, the lowest entrepreneurial activity, the lowest share of the economically active population and also by significant population decline. Territory consists of 10 districts:

Sokolov, Děčín, Chomutov, Litoměřice, Louny, Most, Teplice, Ústí nad Labem, Bruntál and Karviná. The most of these districts can be included among so-called the old industrial areas, where prevails coal mining, metallurgy etc.

References

Czech Statistical Office (2016 a), Share of commuting to work outside the district in total commuting to work in the district, Commuting to work or school, author calculation, available at: https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=profil-uzemi&uzemiprofil=31288&u=__VUZEMI__101__40169# (accessed 13 May 2016).

Czech Statistical Office (2016 b), Share of economically active population of working age (15-64 years), Selected basic indicators by regions and districts (NUTS 3 and LAU 1), author calculation, available at: https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=vystup-objekt-vyhledavani&evo=&str=&udIdent=&zo=N&vyhltext=Tab.34&pvoch=&pvo=ZVCR034&verze=-1&nahled=N&sp=N&nuid=&zs=&skupId=&filtr=G~F_M~F_Z~F_R~F_P~_S~_null_null_&pvokc=&katalog=all&z=T (accessed 10 May 2016).

Decressin, J., Fatás, A. (1995), Regional labor market dynamics in Europe, European Economic Review, 39(9): 1627–1655.

Hančlová, J., Křivý, I., Gottwald, J., Liška, M., Šimek, M., Tvrdík, J., Tvrdý, L. (2002), Modelování a klasifikace regionálních trhů práce, Ostrava: VŠB-TUO.

Hanson, S., Pratt, G. (1992), Dynamic dependencies: A geographic investigation of local labor markets, Economic Geography, 68(4): 373-405.

Holman, R. (2010), Makroekonomie: Středně pokročilý kurz, Praha: C.H. Beck.

Huber, P. (2006), Regional labor market developments in transition, World Bank Policy Research Working Paper No. 3896 (pp. 40), available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=936507 (accessed 11 May 2016).

Ministry of Labour and Social Affairs (MoLSA) (2016 a), Number of applicants per 1 vacancy, Statistical Yearbook of the labor market in the Czech Republic (2013), available at: https://portal.mpsv.cz/sz/stat/stro (accessed 12 May 2016).

Ministry of Labour and Social Affairs (MoLSA) (2016 b), Number of applicants with higher education per 1 vacancy for undergraduates, Statistical Yearbook of the labor market in the Czech Republic (2013), available at: https://portal.mpsv.cz/sz/stat/stro (accessed 12 May 2016).

Ministry of Labour and Social Affairs (MoLSA) (2016 c), Share of long-term unemployed (over 12 moths) in total unemployed, Statistical Yearbook of the labor market in the Czech Republic (2013), available at: https://portal.mpsv.cz/sz/stat/stro (accessed 12 May 2016).

Ministry of Labour and Social Affairs (MoLSA) (2016 d), Share of unemployed since 2005, available at: http://portal.mpsv.cz/ sz/stat/nz/casove_rady (accessed 10 May 2016).

Jurečka, V., Hlaváček, K., Jánošíková, I., Kolcunová, E., Macháček, M., Paličková, I., Spáčilová, L., Wroblovský, T. (2013), Makroekonomie, Praha: Grada.

Jurečka, V., Jánošíková, I., Kolcunová, E., Spáčilová, L. (2009), Mikroekonomie, Ostrava: VŠB-TUO.

Kaufman, L., Rousseeuew, P. J. (2005), Finding groups in data: An introduction to cluster analysis, NewYork: John Wiley & Sons.

Klímek, P. (2008), Shlukovací metody v data miningu, E+M Ekonomie a Management, 11(2): 120-126.

Kutscherauer, A., Fachinelli, H., Sucháček, J., Skokan, K., Hučka, M., Tuleja, P., Tománek, P. (2010), Regionální disparity: disparity v regionálním rozvoji země, jejich pojetí, identifikace a hodnocení, Ostrava: VŠB-TUO.

Layard, R., Nickell, S., Jackman, R. (2005), Unemployment: Macroeconomic performance and the labour market, Oxford: Oxford University Press.

Mach, M. (2001), Makroekonomie II: pro magisterské (inženýrské) studium, Slaný: Melandrium.

Mankiw, N.G. (1999), Zásady ekonomie, Praha: Grada.

Meloun, M., Militký, J. (2002), Kompendium statistického zpracování dat, Praha: Academia.

Pagés, C., Stampini, M. (2009), No education, no good jobs? Evidence on the relation-

ship between education and labor market segmentation, Journal of Comparative Economics, 37(3): 387-401.
Regional Information Systems (RIS) (2016a), Number of registered units per 1000 inhabitants in district, available at: http://www.risy.cz/cs/vyhledavace/statisticka data/detail? Kapitola=4&Ukazatel=9926&RokOd=1995&RokDo=2013&urovenNuts=4&kodNuts=CZ080&MapaRok=2013 (accessed 13 May 2016).
Regional Information Systems (RIS) (2016 b), Total population increase per 1000 mean population, available at: http://www.risy.cz/cs/vyhledavace/statisticka-data/detail?kapitola=3&ukazatel=9946&rokOd=1995&rokDo=2014&urovenNuts=4&nuts2=&nuts3=&nuts4=&orp9= (accessed 11 May 2016).
Romesburg, C.H. (2004), Cluster analysis for researchers, Raleigh: Lulu Press.
Řezanková, H., Húsek, D., Snášel, V. (2007), Shluková analýza dat, Praha: Professional Publishing.
Samuelson, P.A., Nordhaus, D.W. (1991), Ekonomie, Praha: Svoboda.
Tvrdý, L. (2008), Změny na trhu práce a perspektivy vzdělanosti, Ostrava: VŠB-TUO.
Tvrdý, L., Šimek, M., Rievajová, E., Husáková, M., Ivan, I., Machů, M., Konczyna, D., Machová, Z., Mertlová, B., Strážnická, B., Papřoková, A., Holešínská, L. (2007), Trh práce a vzdělanost v regionálním kontextu, Ostrava: VŠB-TUO.
Yonker, S. E. (2016), Geography and the market for CEOs, Management Science, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1625470 (accessed 2 September 2016).
Žák, L. (2004a), Shluková analýza, Automatizace, 47(3): 184-190.
Žák, L. (2004b), Shluková analýza (II), Automatizace, 47(4): 251-254.

Jaroslav Urminský is a PhD student at the Faculty of Economics, VSB – Technical University of Ostrava (Czech Republic). His research and publications focus mainly on the regional economics and development, territorial marketing and labor market. He is the author or co-author of 10 articles. He has been involved in 2 domestic research projects.

Annexes

Annex 1. Descriptive statistics

Descriptive statistics – 4 clusters						
Ward´s method – Clusters		N	Minimum	Maximum	Mean	Std. Deviation
3	Region		1	12	**4,37**	3,011
	Share of unemployed	35	4,6301691820600970	8,9386616667443040	**6,9221603537 32682**	1,2927772869 01994
	Number of applicants per 1 vacancy	35	5,1139489194 49902	28,025969862 135298	**12,545535440 134330**	4,9474895450 26257
	Number of applicants with higher education per 1 vacancy for undergraduates	35	5,7094594594 59460	39,318181818 181820	**19,015479975 114790**	8,3276401066 88307
	Share of long-term unemployed (over 12 moths) in total unemployed	35	22,936415078 621450	46,332131765 307710	**34,491012744 241630**	6,2632964418 01912
	Share of economically active population of working age (15-64 years)	35	67,03617362 069238	72,20516339 308689	**69,75843393 3650250**	1,45659828 3144715
	Share of commuting to work outside the district (another district in the region + other regions + abroad) in total commuting to work in the district	35	5,3704258112 060430	27,574164614 1969600	**11,604952907 997822**	5,0812140299 82537

3	Number of registered units per 1000 inhabitants	35	206	333	**244,66**	27,173
	Total population increase per 1000 mean population	35	-4,9	4,3	**-0,857**	2,3944
	Valid N (listwise)	35				
4.	Region	8	1	10	**2,38**	3,159
	Share of unemployed	8	3,3691024338094384	8,5897541226348950	**6,3532208606 81817**	1,9024755213 77116
	Number of applicants per 1 vacancy	8	7,935979513444302	24,199459093982420	**14,950797013 242154**	6,2343302371 79115
	Number of applicants with higher education per 1 vacancy for undergraduates	8	13,45312500 0000000	91,0000000 00000000	**36,28675829 4142060**	29,18274796 0783143
	Share of long-term unemployed (over 12 moths) in total unemployed	8	16,55080510 8365907	45,3116805 63274890	**33,130090154 085910**	9,00500114 2529348
	Share of economically active population of working age (15-64 years)	8	70,7795002 7872056	73,3628347 8088570	**72,18635851 3768980**	,8391054 49069759
	Share of commuting to work outside the district (another district in the region + other regions + abroad) in total commuting to work in the district	8	19,240170402 9611020	41,11634719 48035450	**28,2342525 52522328**	7,9667933 84258222
	Number of registered units per 1000 inhabitants	8	214	293	**240,63**	28,685
	Total population increase per 1000 mean population	8	4,4	26,4	**11,200**	8,3698
	Valid N (listwise)	8				
2.	Region	10	4	13	**6,50**	3,440
	Share of unemployed	10	8,904018086 2625590	13,4738176403 005220	**11,3257843 23788929**	1,37260686 1112018
	Number of applicants per 1 vacancy	10	15,56902552 2041763	59,061588330 632090	**31,81243409 4749104**	15,1196193 01450559
	Number of applicants with higher education per 1 vacancy for undergraduates	10	8,08474576 2711865	38,500000000 000000	**17,19550353 1556486**	8,473077822 112113
	Share of long-term unemployed (over 12 moths) in total unemployed	10	38,90365873 3949435	50,2061657573 58590	**45,16258073 3227195**	3,543790263 781128
	Share of economically active population of working age (15-64 years)	10	65,27199470 264082	68,272988717 64619	**66,911705037 732100**	,8996240037 20561
	Share of commuting to work outside the district (another district in the region + other regions + abroad) in total commuting to work in the district	10	7,7019359145 527360	18,8936887164 869170	**11,098279815 896374**	3,401704028 934907
	Number of registered units per 1000 inhabitants	10	158	236	**201,70**	20,254
	Total population increase per 1000 mean population	10	-9,9	4,1	**-3,710**	3,7531
	Valid N (listwise)	10				

1.	Region	23	7	13	**10,61**	1,672
	Share of unemployed	23	6,76086488 79048850	11,18186628 77067380	**8,487148052 077078**	1,3913033 39520944
	Number of applicants per 1 vacancy	23	16,6570444 48089223	35,74123788 2177480	**23,9094954298 41850**	5,935522392 565080
	Number of applicants with higher education per 1 vacancy for undergraduates	23	15,6867469 87951807	74,23333333 3333330	**39,943996561 000620**	19,85336668 6533363
	Share of long-term unemployed (over 12 moths) in total unemployed	23	33,0623088 41495530	45,48223030 0628730	**39,30660453 3703080**	3,61930220 0693312
	Share of economically active population of working age (15-64 years)	23	67,0048130 3221030	71,7922841 6403471	**69,60366814 1686530**	1,141570585 091088
	Share of commuting to work outside the district (another district in the region + other regions + abroad) in total commuting to work in the district	23	5,77078126 25977590	20,34930541 57439100	**11,6223851 40705430**	4,0367041 80368968
	Number of registered units per 1000 inhabitants	23	194	262	**215,09**	16,970
	Total population increase per 1000 mean population	23	-7,0	3,6	**-1,535**	2,4740
	Valid N (listwise)	23				

Annex 2. Assigning districts into clusters, Ward´s method

Clusters	Districts
Cluster 1	Hodonín, Přerov, Havlíčkův Brod, Žďár nad Sázavou, Jihlava, Prostějov, Olomouc, Hradec Králové, Blansko, Vyškov, Břeclav, Vsetín, Uherské hradiště, Nový Jičín, Opava, Frýdek-Místek, Svitavy, Šumperk, Znojmo, Kroměříž, Třebíč, Ostrava, Jeseník.
Cluster 2	Bruntál, Karviná, Sokolov, Louny, Litoměřice, Chomutov, Most, Děčín, Teplice, Ústí nad Labem.
Cluster 3	Česká Lípa, Trutnov, Ústí nad Orlicí, Cheb, Jablonec nad Nisou, Karlovy Vary, Liberec, Příbram, Český Krumlov, Tábor, Semily, Strakonice, Tachov, Kutná hora, Mělník, Rakovník, Domažlice, Rychnov nad Kněžnou, Pelhřimov, Jičín, Náchod, Chrudim, Klatovy, Písek, Prachatice, Jindřichův Hradec, Benešov, Rokycany, Mladá Boleslav, Plzeň Jih, České Budějovice, Plzeň-město, Pardubice, Zlín, Brno-město.
Cluster 4	Praha-východ, Praha-západ, Kladno, Kolín, Beroun, Plzeň-sever, Nymburk, Brno-venkov.

Annex 3. One-way Anova

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Region	Between Groups	680,752	3	226,917	29,922	,000
	Within Groups	546,025	72	7,584		
	Total	1226,776	75			
Share of unemployed	Between Groups	177,877	3	59,292	30,127	,000
	Within Groups	141,702	72	1,968		
	Total	319,579	75			
Number of applicants per 1 vacancy	Between Groups	3781,311	3	1260,437	23,052	,000
	Within Groups	3936,804	72	54,678		
	Total	7718,114	75			

Number of applicants with higher education per 1 vacancy for undergraduates	Between Groups	7735,757	3	2578,586	10,527	,000
	Within Groups	17636,889	72	244,957		
	Total	25372,645	75			
Share of long-term unemployed (over 12 moths) in total unemployed	Between Groups	1125,454	3	375,151	11,730	,000
	Within Groups	2302,624	72	31,981		
	Total	3428,078	75			
Share of economically active population of working age (15-64 years)	Between Groups	126,658	3	42,219	26,896	,000
	Within Groups	113,020	72	1,570		
	Total	239,678	75			
Share of commuting to work outside the district (another district in the region + other regions + abroad) in total commuting to work in the district	Between Groups	1998,021	3	666,007	26,868	,000
	Within Groups	1784,759	72	24,788		
	Total	3782,780	75			
Number of registered units per 1000 inhabitants	Between Groups	21535,997	3	7178,666	12,640	,000
	Within Groups	40891,687	72	567,940		
	Total	62427,684	75			
Total population increase per 1000 mean population	Between Groups	1218,897	3	406,299	30,900	,000
	Within Groups	946,727	72	13,149		
	Total	2165,624	75			