
ANALYSIS OF DEVELOPMENT OF LOCAL SELF-GOVERNMENT UNITS IN VOJVODINA

Dragana Tekić¹, Beba Mutavdžić², Tihomir Novaković³, Maja Pokuševski⁴

**Corresponding author E-mail: dragana.tekic@polj.uns.ac.rs*

ARTICLE INFO

Original Article

Received: 05 May 2020

Accepted: 25 May 2020

doi:10.5937/ekoPolj2002431T

UDC 352:502.131.1(497.113)

Keywords:

municipalities, logistic regression, discriminant analysis, prediction

JEL: Q16, M24

ABSTRACT

Discriminant analysis and logistic regressions were applied in this research for the purpose of analyzing the development of autonomous province (AP) Vojvodina local self-government units, which are classified as developed and underdeveloped. The aim of the study is to identify population economic and social characteristics as the one with the most influence on the existence of differences between the observed categories of local self-government units. Based on the results of the discriminatory analysis, number of employed inhabitants per 1,000 inhabitants and number of highly educated inhabitants per 1000 inhabitants were found to have the greatest influence on the development of the local self-government unit, while based on logistic regression results, number of employed inhabitants per 1000 inhabitants and natural increase are the most influential factors. Both models have good data classification power, the discriminant analysis model successfully classifies 90.9% of all cases, and the logistic regression model successfully classifies 88.6% of cases.

© 2020 EA. All rights reserved.

Introduction

The effects of the globalization are manifested not only at the national level, but also at the level of mesoregions or micro-regions, which increase the importance of territorial units. This stems from the fact that local and regional development responsibilities

- 1 Dragana Tekić, MAgrEC, Junior Researcher, University of Novi Sad, Faculty of Agriculture, Trg D. Obradovića 8, 21000 Novi Sad, Serbia, Phone: +381 (21) 4853 380, e-mail: dragana.tekic@polj.uns.ac.rs, ORCID ID (<https://orcid.org/0000-0002-1924-6196>)
- 2 Beba Mutavdžić, PhD, Professor, University of Novi Sad, Faculty of Agriculture, Trg D. Obradovića 8, 21000 Novi Sad, Serbia, Phone: +381 (21) 4853 382, e-mail: beba.mutavdzic@polj.uns.ac.rs, ORCID ID (<https://orcid.org/0000-0002-7631-0465>)
- 3 Tihomir Novaković, MAgrEC, Teaching Assistant, University of Novi Sad, Faculty of Agriculture, Trg D. Obradovića 8, 21000 Novi Sad, Serbia, Phone: +381 (21) 4853 380, e-mail: tihomir.novakovic@polj.uns.ac.rs, ORCID ID (<https://orcid.org/0000-0002-8405-3403>)
- 4 Maja Pokuševski, Master student, University of Novi Sad, University center of applied statistics, Železnička 19, 21000 Novi Sad, Serbia, Phone: +381 643749 893, e-mail: majapokusevski@gmail.com, ORCID ID (<https://orcid.org/0000-0003-1906-594X>)

and competencies are delegated to the regional institutions (Liptáková, Rigová, 2020). Assessing regional and, as well as, local development is a methodologically challenging and politically relevant issue. The development of a region depends on the development level of the local governments in that region. Through local economic development, the economic capacity of the local area is developed to create a basis for economic progress and quality of life for the whole society. Local economic development integrates regional and development policy, as well as all other policies, with the aim of faster development of local communities (Glavaš-Trbić, et al. 2008). Local economic development is composite and complex area that, in addition to economic development policy including agriculture, also incorporates other divisional, structural and social policies, local infrastructural development policy, as an indispensable ambience for local economic development, as well as all sorts of civic initiatives contributing to local communities' improvement (Kačar, et al. 2016).

The aim of this research is to determine the influence of various factors on the development of the observed units in local self-government (municipalities) in the Vojvodina region by applying discriminant analysis and logistic regression, as statistical methods suitable for the categorical data analysis. Specifically, the factors that are expected to have an impact on the development of a particular municipality are: population density - population per km² (PD), number of employed inhabitants per 1,000 inhabitants (EM), number of highly educated inhabitants per 1,000 inhabitants (ED), natural increase (NI) and investment in new capacities (IN).

Materials and methods

The classification of local self-government units into developed and underdeveloped ones was carried out based on the "Decree on the establishment of a single list of development of the region and local self-governemnt units for 2014". Regions and local self-government units, which are classified into the first, second, third and fourth groups and devastated areas based on data from the authority responsible for statistics and finance. ("Sl. glasnik RS", No. 104/2014). The classification of regions and local self-government units into specific groups was done based on the gross domestic product per capita value in the region or local self-government unit, relative to the national average. For the purposes of this research, local self-government units are classified as developed (first and second group), development rate is over 80% of the national average and underdeveloped (third and fourth group), development rate is below 80% of the national average.

For statistical analysis of selected factors of development of local self-government units (municipalities), two statistical methods were applied: discriminant analysis and binary logistic regression. Discriminant analysis (DA) and logistic regression (LR) are widely used multivariate statistical methods for analyzing data with categorical outcome variables (Pohar, et al. 2004). The difference between these two methods is that the discriminant analysis implies certain assumptions that must be respected for its application, above all the normality of the data, while the logistic regression model is not based on any assumptions.

Discriminant analysis

Discriminant analysis is a multivariate technique which focuses on association between categorical dependent variables and multiple independent variables (Ahsan ul Haq et al., 2015).

Discriminant analysis is a parametric model of multivariate analysis that is based on the following assumptions:

- 1) there is no high correlation of explanatory variables,
- 2) variance and covariance of individual groups explanatory variables pairs are equal (homogeneous) and
- 3) explanatory variables have a normal distribution. (Sokolovska et al., 2014).

Kolmogorov-Smirnov and Shapiro-Wilks normality tests, Leven variance homogeneity test and Brown-Forsythe arithmetic mean group test were used to test the assumptions for discriminant analysis. The homogeneity of the group covariance matrices was checked using Box's M statistics.

Some authors define that the variate for a discriminant analysis, also known as the discriminant function, is derived from an equation much like that seen in multiple regression. It takes the following form (Hair et al., 2006):

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + \dots + W_nX_{nk} [1]$$

Z_{jk} = discriminant Z score of discriminant function j for object k

a = intercept

W_i = discriminant weight for independent variable i

X_{ik} = independent variable i for object k

Wilk's λ -test was used to interpret the obtained discriminant function, which is of the differences among group means of independent variables, was used to ascertain the level of significance for each group predictor. To estimate the degree of deviation influence, the standardized canonical discriminant function was applied (Heil, Schmidhalter, 2014).

Logistic regression

Logistic regression model represents a statistical method for predicting the outcome of categorical dependent variable based on one or more independent variables that are called predictors. When observed outcome for dependent variable has two possible options, model is called binary logistic regression model (Kovljenić, Savić 2017).

The following form of regression is used for this purpose:

$$\pi(x) = \frac{e^{\alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}}{1 + e^{\alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}} \quad [2]$$

Where $\pi(x)$ represents the expected value of Y for a given value of X , while the parameters α i $\beta_1, 2, \dots, k$ correspond to the parameters α i $\beta_1, 2, \dots, k$ from the linear regression model and represent the average initial level of the dependent variable and coefficients regressions showing the average change in logit per unit of change independently variable. The logistic regression function thus obtained is nonlinear and can be linearized by logit transformation.

If the logistic regression function is linearized, we get the following form:

$$\ln\left(\frac{\pi}{1 - \pi}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad [3]$$

The resulting equality is called logit and it is linear with the parameters β_i , $i = 1 \dots k$. It can be observed that π belongs to the interval $[0, 1]$, while the logit value ranges from $(-\infty, +\infty)$, so it can be said that the logit function is the best choice for displaying this function (Chatterjee, Ali, 2006). The Wald statistic test is usually used in which β is estimated using the maximum likelihood estimator (Basu et al., 2017).

The overall assessment of the model to fit the data can be examined using the Hosmer-Lemeshow test, as well as the classification matrix provided by the SPSS software package used in the data processing. One of the most commonly used indicators of model quality is Cox and Snell and Nagelkerke pseudo R^2 . Although values of pseudo R^2 indices typically range from zero to unity, values for some indices can exceed 1.0 (Walker, Smith, 2016).

The choice of variables is conditioned by many factors, the most important of which are the availability of data and the requirements set by the applied statistical methods. The survey is based on data about the development of AP Vojvodina local self-government units from the "Municipalities and regions" (Opštine i regioni) for the period 2013-2018. The SPSS software package was used for statistical data processing.

Results and Discussion

From the Table 1 it can be seen that out of 45 local self-government units in the territory of AP Vojvodina, 26 have the status of developed units of local self-government, while the other 19 have the status of underdeveloped local self-government units. In regional terms, the average number of inhabitants per km^2 is 82, the smallest number of inhabitants per km^2 is in the municipality of Sečanj 23, and the largest in Novi Sad with 528 inhabitants per km^2 .

In terms of employment, the average at the regional level is 232 employees per 1,000 inhabitants, which shows a low employment rate. The city of Novi Sad has the highest employment rate with 400 employees per 1,000 inhabitants, while the municipality of Opovo has the lowest employment rate with 126 employees per 1,000 inhabitants.

The average number of university graduates per 1,000 inhabitants in the territory of Vojvodina is 91, the lowest number of higher educations is in the municipality of Žabalj, while the highest number is those with higher education in Novi Sad.

A negative natural increase rate is present in almost all municipalities in the territory of Vojvodina, only the city of Novi Sad stands out with a positive natural growth rate of 0.8 ppm. Investments in new capacities are presented in absolute amount. The average investment in the observed period amounts to RSD 2,737,636.91. High values of coefficients of variation indicate that there are significant differences between the observed municipalities. The highest variability is observed with the variable investment, which is expected given the variation range.

Table 1. Descriptive statistics

Variable	Mean	Minimum	Maximum	Coefficient of variation (%)
PD	82	23	528	95.68
EM	232	126	400	25.71
ED	91	49	219	34.26
NI	-7.38	-12.2	0.8	34.09
IN	2,737,636.91	0	34,434,118.00	220.26

Source: Authors calculation

Firstly, the assumptions for applying discriminatory analysis were tested. The first assumption refers to the collinearity of the variables, and for the purpose of testing the collinearity of the variables, a correlation matrix was used within the groups to show the correlation between the variables (Table 2). Table 2 shows that the highest values of correlation coefficients are visible in the correlation between PD and IN ($r = 0.665$), followed by PD and ED ($r = 0.593$) and PD and NI ($r = 0.546$).

Table 2. Correlation matrix

Variable	PD	EM	ED	NI	IN
PD	1.000	0.461	0.593	0.546	0.665
EM		1.000	0.458	0.415	0.541
ED			1.000	0.511	0.537
NI				1.000	0.416
IN					1.000

Source: Authors calculation

Testing the homogeneity of variance of individual variables between groups was performed using the Leven test for testing the homogeneity and the Brown-Forsythe test of arithmetic means of groups equality (Table 3).

Table 3. Results of Levene's and Brown-Forsythe tests

Variable	Levene's statistics	Sig.	Brown-Forsythe statistics	Sig.
PD	5.611	0.022	12.172	0.002
EM	5.174	0.028	49.496	0.001
ED	6.770	0.013	35.767	0.001
NI	0.454	0.504	9.935	0.003
IN	7.760	0.008	7.261	0.012

Source: Authors calculation

The Leven test results for all variables except for the natural increase variable show the heterogeneity of variance. As the Brown Forsythe test was applied to test groups in the case of heterogeneous variance, the data presented in Table 3. shows statistically significant group mean.

The application of discriminant analysis assumes the existence of group covariance matrices homogeneity, which is usually checked in Box's M statistics in multivariate analysis (Table 4). This test statistical significance may be due to the deviation of the data from the normal distribution, not to the inequality of the metrics covariance.

The results presented in Table 4 show that complete agreement with the multidimensional normal distribution was not reached.

Table 4. Results of Box's M statistics

Box's M		136,478
	Approx.	7,860
F	df1	15
	df2	5338,556
	Sig.	.000

Source: Authors calculation

The last assumption of discriminant analysis concerns the normality and linearity of the original data. Apart from the fact that all variables, except for NI and EM, show deviations from the normal distribution, the original data is burdened with many non-standard observations. Since the original set of variables did not achieve complete agreement with the normal distribution, logarithmic transformation of the data was applied.

The transformation achieved not only better agreement of the transformed data distribution with normal distribution, but also a reduction in the number of non-standard observations, which gives the analysis better opportunities to more accurately extract discriminatory functions.

Results of discriminant analysis

From the results shown in Table 5 a single canonical discriminant function was isolated.

Table 5. Results of discriminant function

Function	Eigenvalue	Canonical Correlation	Wilks' Lambda	Chi-square	Sig.
1	1.523	0.777	0.396	36.559	0.001

Source: Authors calculation

The eigenvalue indicates the relative discriminant power of the discriminant function, the higher eigenvalue means that the more variance in the dependent variable is explained by the given function. The canonical correlation is 0.777; it represents the quadratic root of the relation between the intergroup and the total sum of squares.

The significance of the isolated discriminant function was tested via Wilks' lambda = 0.396 and, for $\chi^2 = 36,559$ and $df = 5$, confirmed at $p = 0.000$, which, together with the value of the canonical correlation coefficient, show that discriminative function is significant (Table 5).

In the Structure matrix table (Table 6) variables are ordered by absolute values of correlations with the discriminant function.

Table 6. Structure matrix

Variable	Function
logED	0.792
logEM	0.776
logPD	0.653
logIN	0.647
logNI	0.363

Source: Authors calculation

The largest contribution to the discriminatory function structure were made by variables: ED (0.792), followed by EM (0.776). The smallest contribution to the discriminant function structure had the variable NI (0.363). Although significantly different, the values of all coefficients are statistically significant.

Table 7. Discriminant function coefficients

Variable	Standardized discriminant function coefficients	Discriminant function coefficients
logPD	0.297	1.446
logEM	0.465	0.011
logED	0.433	4.510
logNI	-0.126	-0.054
logIN	0.229	0.438
Constant		-16.884

Source: Authors calculation

The discriminant function standardized canonical coefficients (Table 7) represent measure of the selected independent variables relative influence, the higher value of the coefficients corresponds to the greater discriminative ability and means that the groups differ in that variable. The independent variable with the most discriminatory power is EM, followed by ED, while the other three independent variables were less successful as predictors. Canonical discriminant function coefficients represent the coefficients of final canonical discriminant function (Table 7).

Based on the calculated coefficients, the discriminant function takes the following form:

$$Z = -16,884 + 1,446\log PD + 0,011\log EM + 4,510\log ED - 0,054\log NI + 0,438\log IN$$

After discriminant function was calculated, the intersection point was determined based on the centroids in each group.

Table 8. Function at group centroids

Development of local self-government units	Function
Underdeveloped	-1.449
Developed	1.003

Source: Authors calculation

The discriminant function intersection point is weighted average between the centroids in each of the distributions. The optimum cross-sectional limit recorded is 1.003. This value classifies municipalities according to their discriminatory result, i.e. municipalities where function value is below 1.003 belong to the group of underdeveloped municipalities, while municipalities with discriminatory grades above this value belong to the group of developed municipalities (Table 8).

Results of logistic regression

The stepwise method was used to select variables in the regression analysis. The selection of variables is conducted in four steps, from which only the results of the fourth step will be described.

The performance of the model was tested using the Omnibus coefficient test, called also as “goodness of fit” because it shows how well the model predicts results.

Table 9. Omnibus tests of model coefficients

Step	Chi-square	df	Sig.
Step 4 Block	39.676	2	0.001
Model	39.676	2	0.001

Source: Authors calculation

The Omnibus test (Table 9) found that there was a statistically significant difference between the models containing the selected independently variable and the one containing no independent variable (Sig. <0.05). The same conclusion can be drawn from the data presented in the following table.

Table 10. Hosmer and Lemeshow test results

Step	Chi-square	Df	Sig.
4	4.054	8	0.852

Source: Authors calculation

In the case of the Hosmer and Lemeshow test the indicator of poor prediction is a Sig. value of less than 0.05. In the analyzed example, the value for the Hosmer and Lemeshow test is 4.054 with a significance of 0.852, which leads to the conclusion that the model can be the basis for the prediction.

Model fit was estimated using Cox and Snell's and Nagelkerke Pseudo R-Square coefficients. The values of these coefficients indicate how accurately the model explains the analyzed data set.

Table 11. Model summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
4	19.859	0.594	0.801

Source: Authors calculation

The third and fourth columns (Table 11) show the values of pseudo coefficients, the values of these two indicators are 0.594 and 0.801, which indicate that the model with the given set of variables is well fitted to the data.

Table 12 presents information about contribution or importance of each predictor variable. The contribution of predictor variables were valuated based on the results of the Wald test.

Table 12. Variables in the equation

Step	Variables	B	S.E.	Wald	df	Sig.	Exp(B)
Step 4	EM	0.081	0.028	8.386	1	0.004	1.084
	NI	0.816	0.378	4.649	1	0.031	2.261
	Constant	-10.506	4.597	5.224	1	0.022	0.000

Source: Authors calculation

Based on the Wald test Sig. value presented in the Table 12, it can be concluded that only two, of the observed predictor variables, have statistical significance.

In this analysis, the main factors that influence whether a municipality will be developed are EM (Sig = 0.0004) and NI (Sig = 0.031), while other factors did not significantly contribute to the model predictive capabilities.

Based on the predictor variables calculated coefficients the logistic regression model equation is calculated, and it takes the following form:

$$Y = -10.506 + 0.081EM + 0.816NI$$

An area under the rock curve (AUC) was calculated, for the purpose of additional analysis on the degree of the prediction agreement with the data.

Table 13. Area under the curve

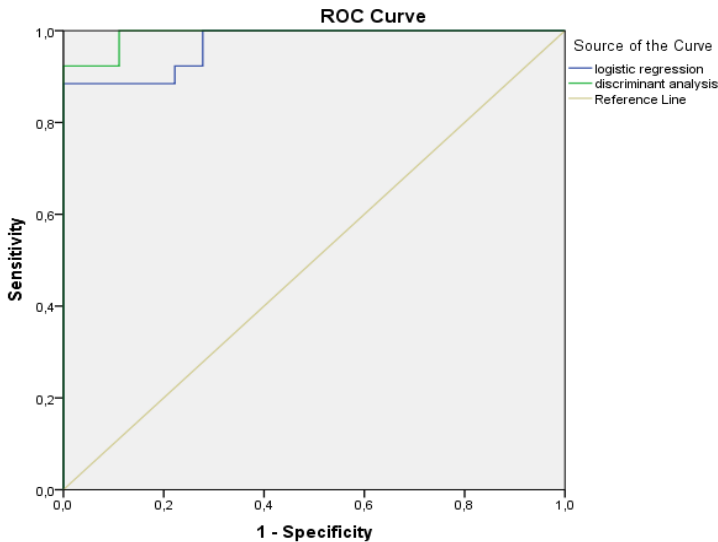
Variable	Area	Std. Error	Asymptotic Sig.	Asymptotic 95% confidence interval	
Discriminant analysis	0.991	0.021	0.000	0.929	1.000
Logistic regression	0.970	0.009	0.000	0.974	1.000

Source: Authors calculation

The results (Table 13) show that the AUC for the logistic regression model is 0.970, while the AUC for the discriminant analysis model is 0.991. The area mentioned here speaks of extraordinary separation.

The ROC curve to which the above analyzes refer is shown in Figure 1.

Figure 1. ROC curve



Source: Authors calculation

Data presented in the Table 14 show how accurately the model predicts categories of dependent variables. The discriminant analysis model successfully classifies 90.9% of all cases, while the logistic regression model successfully classifies 88.6% of all cases.

Table 14. Classification table

Development	Discriminant analysis		Logistic regression	
	Underdeveloped	Developed	Underdeveloped	Developed
Underdeveloped	16	2	16	2
Developed	3	24	3	23
Total (%)	90,9		88,6	

Source: Authors calculation

Based on the values in the classification table, it is possible to determine the sensitivity of the model (Table 15).

Table 15. Comparison of models

Discriminant analysis			Logistic regression		
Sensitivity (%)	Specificity (%)	AUC (%)	Sensitivity (%)	Specificity (%)	AUC (%)
92,31	84,21	99	92	84,21	97

Source: Authors calculation

It can be noted (Table 15) that logistic regression and discriminant analysis models have successfully classified approximately the same percentage of cases. However, based on the AUC values, it can be concluded that the discriminant analysis model slightly exceeds the logistic regression model.

Conclusion

This paper compares two methods: discriminant analysis and logistic regression to assess the impact of five variables on the likelihood of a local government unit being classified as developed or underdeveloped. Variables that were assumed to have an impact on the development of municipalities are: population density - number of inhabitants per km², number of employees per 1,000 inhabitants, number of highly educated inhabitants per 1,000 inhabitants, natural increase and investments in new capacities. Out of 45 municipalities in AP Vojvodina, 26 municipalities belong to the group of developed, while 19 municipalities have the status of underdeveloped municipalities. After testing the assumptions for the application of discriminant analysis, the discriminant function was calculated. The discriminatory analysis results showed that the most important factors influencing the municipality classification are number of employees per 1,000 inhabitants and the number of higher education inhabitants per 1,000 inhabitants. The significance of the discriminant function was confirmed by the Wilks' lambda test and the canonical correlation coefficient value. The logistic regression results showed that number of employees per 1,000 inhabitants and natural increase are the most important predictors. The model evaluation was performed by measuring the overall classification accuracy, sensitivity and specificity as well as by examining the area under the ROC curve (AUC). The results show that both models have good classification power. The discriminant analysis model successfully classified 90.9% of all cases, while the logistic regression model 88.6% of all cases. When considering the percentages of sensitivity,

specificity and AUC, it can be observed that the differences between the two models are insignificant, but in the specific example the discriminant analysis model gave better results and should be used as a basis for prediction.

Conflict of interests

The authors declare no conflict of interest.

References

1. Ahsan ul Haq M., Irum Sajjad D. & Qura-tul-ain. (2015): Performance comparison of classification techniques, artificial neural network, discriminant analysis & logistic regression, *Science International*, 27(3), 1803-1807.
2. Basu A., Ghosh A., Mandal A., Mart'ın N. & Pardo L. (2017), A Wald-type test statistic for testing linear hypothesis in logistic regression models based on minimum density power divergence estimator, *Electronic Journal of Statistics*, 11(2), 741-2772. DOI: 10.1214/17-EJS1295
3. Chatterjee S. & Ali S. H. (2006): *Regression analysis by example*, Fourth edition, John Wiley & Sons, New York.
4. Glavaš-Trbić D., Pejanović R. & Maksimović G. (2008): Rural Development and Local Economic Development of Serbia, *Agroeconomics*, Department of Agricultural Economics and Rural Sociology, Faculty of Agriculture, Novi Sad, Serbia, 47 (48), 80- 91. [In Serbian: Glavaš-Trbić D., Pejanović R. & Maksimović G. (2008): Ruralni razvoj i lokalni ekonomski razvoj Srbije, *Agroekonomika*, Departman za ekonomiku poljoprivrede i sociologiju sela, Poljoprivredni fakultet, Novi Sad, Srbija, 47 (48), 80- 91].
5. Hair J., Black W., Babin B., Anderson R., Tatham R. (2006): *Multivariate data analysis*, Pearson Prentice Hall, New Jersey.
6. Heil K. & Schmidhalter U. (2014): Using discriminant analysis and logistic regression in mapping quaternary sediments, *Mathematical Geosciences*, 46(3),361-376. DOI: 10.1007/s11004-013-9486-x
7. Kačar B., Curić J. & Ikić S. (2016): Local economic development in theories of regional economies and rural studies, *Economics of Agriculture*, 63(1), 231-246. DOI:10.5937/ekoPolj1601231K
8. Kovljenić M. & Savić M. (2017): Factors influencing meat and fish consumption in Serbia households: Evidence from SILC database, *Economics of Agriculture*, 64(3), 945-956. DOI: 10.5937/ekoPolj1703945K
9. Liptáková K. & Rigová Z. (2020): Possibilities of Slovak municipalities to participate in regional development in context of globalization, *The 19th International Scientific Conference Globalization and its Socio-Economic Consequences -Sustainability in the Global-Knowledge Economy*, Rajecke Teplice, Slovakia, 74 (1),1-8. DOI:10.1051/shsconf/20207405013

10. Pohar M., Balas M. & Turk S. (2004): Comparison of Logistic Regression and Linear Discriminant Analysis: A Simulation Study, *Metodološki zvezki*, 1(1), 143-161.
11. Sokolovska V., Nikolić- Đorić E. & Žolt L. (2014): Regional differences in the Republic of Serbia, Regions and regionalization, Faculty of Philosophy, University of Novi Sad, Serbia, no. 3, 9-22. [In Serbian: Sokolovska V., Nikolić- Đorić E., Žolt L. (2014): Regionalne razlike u Republici Srbiji, Regioni i regionalizacija, Filozofski fakultet, Univerzitet u Novom Sadu, Srbija, no. 3, 9-22].
12. The Statistical Office of the Republic of Serbia, Municipalities and regions.[in Serbian: Republički zavod za statistiku, Opštine i regioni]. Retrieved from www.stat.gov.rs, (February 16,2020).
13. The Official Gazette Republic of Serbia, Decree on the establishment of a single list of development of the region and local self-governemnt units for 2014, No. 104/2014. [in Sebian: Sl. glasnik RS, Uredba o utvrđivanju jedinstvene liste razvijenosti regionai jedinica lokalne samouprave za 2014. Godinu, No. 104/2014].
14. Walker D. & Smith T. (2016): JMASM Algorithms and code nine pseudo R indices for binary logistic regression models, *Journal of Modern Applied Statistical Methods*, 15(1), 848-854.