

## DETERMINATION OF THE LEVEL OF SUSTAINABLE DEVELOPMENT OF THE CITIES – A PROPOSAL FOR A METHOD OF CLASSIFYING OBJECTS BASED ON NATURAL BREAKS

Krzysztof Rząsa<sup>1</sup>, Mateusz Ciski<sup>2</sup>

<sup>1</sup> ORCID: 0000-0001-5215-7630

<sup>2</sup> ORCID: 0000-0002-4919-9570

<sup>1,2</sup> University of Warmia and Mazury in Olsztyn

Prawocheńskiego street 15, 10-720 Olsztyn, Poland

### ABSTRACT

**Motives:** The problem of measuring the level of sustainable development is a subject addressed by many authors in their research.

**Aim:** In this article the Authors proposed a new method of classifying objects based on Jenks' Natural Breaks to measure the level of sustainable development. The analysis was carried out on the basis of the data obtained from Statistics Poland. An important element of the research was the development of the process of selection and rejection of input data on the basis of a variety of statistical indicators. This resulted in a set of data which, on the one hand, is statistically justified and, on the other, describes the examined phenomenon in a comprehensive way.

**Results:** The research objects were 66 Polish district cities; Authors obtained a ranking of cities in terms of its Sustainable Development Level. The authors decided to verify the correlation of the results obtained from a proposed method of classifying objects based on natural breaks, with those from the chosen taxonomic method (Hellwig's method) and the Classic Ranking. The fact of receiving highly correlated results confirms the validity and reliability of the proposed method.

**Keywords:** sustainable development, Jenks' Natural Breaks, Hellwig's method, Classic Ranking method.

### INTRODUCTION

Sustainable development is usually defined, in the words of The Brundtland Report "Our Common Future", as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [World Commission on Environment and Development, 1987]. The report highlighted three fundamental components of sustainable development: environmental

protection, economic growth, and social equity. The definition of sustainable development can be found in Polish legislation – it is defined in the Environmental Protection Act [Act of 27 April 2001 Environmental Protection Law, 2001] as the socio-economic development integrating political, economic and social actions, balanced with environmental protection and permanence of basic natural processes, in order to ensure the possibility of satisfying the basic

✉ [krzysztof.rzasa@uwm.edu.pl](mailto:krzysztof.rzasa@uwm.edu.pl), ✉ [mateusz.ciski@uwm.edu.pl](mailto:mateusz.ciski@uwm.edu.pl)

needs of communities or individual citizens in both the present and future generations.

Sustainable development is a sort of compromise between social, economic and environmental goals, representing the well-being of present and future generations.

The social aspect is identified with education and acquiring the ability to solve major social problems as well as participation in the development processes of the whole system. Within the social domain, the following areas are analyzed: access to labor market, criminality, demographic changes, determinants of health, old-age income adequacy, poverty and living conditions, public health, road accidents, and sustainable consumption patterns. The economic aspect of sustainable development, on the other hand, means not only meeting present needs, but also securing the resources necessary to meet the needs of future generations (within this domain there are the areas of: intellectual and social capital, economic development, and transport). Whereas the environmental aspect means establishing the limits of the natural system for human activities and not exceeding them. Within the environmental domain, the following areas are analyzed: climate change, energy, land use, biodiversity, and waste management.

The approach of the Polish authorities is consistent with the EU approach. The definition included in the Environmental Protection Act corresponds to the assumptions adopted for the European Union area. Similar sets of indicators are used to measure Sustainable Development both in Poland and in other EU countries, which is in line with current EU guidelines [European Commission, 2016].

Sustainable development is nowadays one of the main component used in the spatial development [Alkhalidi et al., 2018, Antonopoulos, 2018, Bell & Morse, 2018, Ciski et al., 2019]. All existing definitions of sustainable development indicate a very broad scope of this concept. In scientific research, many authors propose different ways of measuring the level of sustainable development, but due to the broad nature of the concept, these methods are very diverse and are based on different methodological approaches [Atkinson et al., 1997, Dasgupta, 2007, Moran et al.,

2008, Moreno Pires et al., 2014, Mortensen, 2013, Nourry, 2008, Pearce et al., 1996; Tanguay et al., 2010]. In this article, the Authors decided to propose a method of classifying objects based on natural breaks for this purpose.

Classification methods have been applied to many applications in various fields of science [Aly, 2005, Sharma et al., 2016, Tharwat, 2018]. Data classification is a complex process that may be affected by many factors [Lu & Weng, 2007]. Different techniques are used for classification, including probabilistic methods, decision trees, rule-based methods, instance-based methods, support vector machine methods, and neural networks [Aggarwal, 2014]. There are qualitative and quantitative research methods among them [Toloie-Eshlaghy et al., 2011]. Jenks natural breaks optimization method is one of the most popular data clustering methods used to classify objects [Chen et al., 2013, Khamis et al., 2018], it is used to determine the best distribution of values in different classes. Jenks' algorithm uses an iterative approach to find the best groupings of numbers based on how close they are together (based on variance from the group's mean) while also trying to ensure the different groupings are as distinct as possible (by maximizing the group's variance between groups). Thus, the Jenks optimization method aims to minimize the average deviation of each class from the average class; reducing variance within classes and maximizing variance between classes. Jenks optimization method's algorithm can be presented in the following repeated steps:

1. Compute the sum of squared deviations from the class means (SDCM).
2. Compute the sum of squared deviations from the array mean (SDAM).
3. After inspecting each SDCM, a decision is then made to move one unit from a class with a larger SDCM to an adjacent class with a lower SDCM.

New class deviations are then computed, and the process is repeated until the sum of the within class deviations reaches a minimal value [Jenks, 1967]. Taxonomic methods are used to describe and classify complex socio-economic phenomena expressed by a large set of variables [Bąk, 2016, Podogrodzka,

2011, Prus & Król, 2017, Senetra & Szarek-Iwaniuk, 2019]. Taxonomic methods can be divided into hierarchical and non-hierarchical, agglomeration and divisional, area-based, and optimization [Grabiński et al., 1989]. Among the hierarchical methods, linear ordering methods are an important group; these methods determine the linear hierarchy of objects based on the distance from the computed so-called “development pattern”. All methods allow for carrying out classifications on a multidimensional set of variables, but they differ mainly in computational algorithms. Hierarchical methods allow for separating the full hierarchy, i.e. the focus on the higher level includes detachable of the lower levels forming cluster structures. In non-hierarchical methods it is not possible to present the process of cluster formation as a structure. In agglomeration methods, each of the research objects is initially a separate cluster, and the process ends with the merger of all objects into one group. Divisional methods treat a set of objects as one group, successive divisions lead in the final effect to the separation of the number of groups equal to the number of objects. In area-based methods, hyper-area is divided into separable areas, while in optimization methods, using different criteria, there are to further approximations of the division of the collection [Grabiński et al., 1989, Kuciński, 2015, Prus & Król, 2017]. To perform a multidimensional spatial analysis of phenomena, a classification procedure can be applied, which consists of separating homogeneous subsets in a set of multi-characteristic objects. The classification is based on the assumption that objects belonging to the same cluster will be similar (homogeneous), while those belonging to different clusters will be heterogeneous [Hellwig, 1981]. Separation of homogeneous groups of objects makes it possible to carry out in-depth analyses within those groups taking into account a larger number of variables. This approach allows for a better understanding of the factors determining the level and structure of the examined phenomenon, and consequently for a more accurate assessment of the reality and identification of possible causes differentiating the compared objects [Gorzela, 1981].

The choice of the variables (features) is one of the most important steps in research; variables used in applied research should be selected carefully [Tarka, 2010]. Diagnostic features to classify objects should have specific properties. This was written about as early as 1957, by one of the precursors of taxonomic methods, Prof. Fierich. According to [Fierich, 1957], variables ought to:

- include the most important properties of the analyzed phenomena, therefore only the necessary properties should be taken into account;
- be simply and logically connected;
- be clearly and strictly defined and directly or indirectly measurable and expressible by absolute (similarity of magnitude) or relative (similarity of structure) quantities;
- have high spatial variability in the set of tested objects and cannot be easily influenced by the environment;
- be independent of each other, but related to characteristics not included in the study;
- have high coefficients of variation within the initial community and, within groups, as low as possible.

The requirements for variables were also similarly defined in other subsequent studies. For example, according to [Gorzela, 1979], a good set of variables are those that:

- are strictly defined;
- are unambiguous;
- represent as accurately as possible the phenomena and processes falling within the scope of research;
- have a high information content;
- are uncorrelated.

Based on the research of its predecessors, the problem of the properties of diagnostic features has been described most extensively by [Grabiński et al., 1989] and is still the basis for this type of research. According to this, the diagnostic features should:

- capture the most important properties of the analyzed phenomena and represent it accurately;
- be simple, clear, and precise;
- be logically connected;
- contain a high content of information;

- be directly or indirectly measurable, which boils down to the existence of reliable and easily accessible statistical data;
- be expressed in natural rather than valuable units, rather than in absolute terms;
- be characterized by high spatial variability;
- not be highly correlated;
- be highly correlated with undiagnostic and synthetic variables;
- enable mutual control (through knowledge of the statistical and substantive relationships between individual variables);
- not to describe specific phenomena and processes;
- be characterized by the consistency of proportions between the number of variables characterizing a given aspect of the examined phenomena and their substantive significance.

The main aim of the article was to propose a method of classifying the level of sustainable development on the basis of natural breaks. This method was used to determine the level of sustainable development of 66 Polish cities. In order to verify the usefulness of the proposed method, the results were compared with the results obtained using the Classic Ranking method and the selected taxonomic method – Hellwig’s method.

## MATERIALS AND METHODS

Research carried out in this study can be divided into three stages:

Stage I: the selection of diagnostic variables.

Stage II: determination of the sustainable development level using the proposed method.

Stage III: determination of the sustainable development level using Classic Ranking and Hellwig’s methods.

### Selection of diagnostic variables

The selection of diagnostic variables used to determine the level of sustainable development of cities was primarily based on a literature analysis and datasets of indicators available on the Internet to assess sustainable development [Azapagic & Perdan,

2000, Bossel, 1999, Drastichová, 2017, European Commission, 2015, Hák et al., 2016, Klopp & Petretta, 2017, Mori & Christodoulou, 2012, Morton et al., 2017; Schleicher-Tappeser, 2018, Shen et al., 2011, Spangenberg, 2015, United Nations, 2007, 2015]. The results of sustainability studies carried out by different researchers for different objects, such as countries, administrative parts of countries, etc., were also analyzed [Bąk & Cheba, 2018, Czermińska, 2002, Koszel & Bartkowiak, 2018, Moran et al., 2008, Mortensen, 2013, Stafford-Smith et al., 2017]. An equally important element of determining the selection of diagnostic variables was its availability and reliability. Therefore, the authors decided to choose a set of indicators published in the Sustainable Development Indicators (SDI) service [Sustainable Development Indicators, 2017]. All data comes directly from Statistics Poland, the central office of government administration in Poland, which collects and shares statistical information. The SDI service is used to disseminate and demonstrate indicators for monitoring sustainable development at national, regional, provincial, and district levels. Currently, it is difficult to obtain detailed data for Poland’s territorial division units – data in the new Sustainable Development Goals panel in Statistics Poland adapted to United Nations Sustainable Development Knowledge Platform [United Nations, 2015] are published only at the national level. In order to examine the level of sustainable development of Poland’s largest cities, the Authors have used data from 2016, i.e. the last year for which the data were published in the “Sustainable Development Indicators” service.

The input data for analysis is divided into four levels, taking into account the data hierarchy and based on the detail of the data:

1. Domains – the first level, grouping areas, and indicators; these are the four main branches of statistical data that make up the idea of sustainable development: Social, Economic, Environmental, and Institutional-political.

2. Areas – the second level, grouping indicators; areas consist of more specific data, for example Demographic changes, Poverty and living conditions, Criminology, Openness and participation, etc.

3. Indicators – statistical indicators, embedded in the idea of sustainable development; the indicators are, for example Demographic dependency ratio, Average useful floor area of dwelling per capita, Ascertained by Police crimes total per 1,000 population, etc.

4. Dimensions – indicators are additionally divided into different dimensions; most indicators are described by two or more dimensions, for example: “Length of bicycle lane” indicator is described with “per 10 thousand km<sup>2</sup>” and “per 10 thousand population” dimensions.

The input database consisted of 119 indicators, divided into four domains with 21 areas. The following four tables prepared separately for each domain, list all indicators. The first column in the tables is the number of the indicator – from X<sub>1</sub> to X<sub>119</sub>; each row means one indicator. The second column contains the thematic groups of the indicators. Many indicators vary in dimension while sharing the same group; in such cases, to make tables more readable, the indicator group cells have been merged. The third column contains the dimension and information about the units of measurement of indicators.

The first table contains information about the social domain. There are 59 indicators in this domain, nearly 50% of the total input data. The indicators included in this domain describe the standard of living of local communities in the analyzed cities; it concerns demographic changes, health and living conditions,

labor market, but also crime and road accidents. Table 1 shows the list of social domain indicators.

The second table gathers information on the economic domain describing the economic situation of the analyzed cities, including: investment expenditures in enterprises, expenditures on innovative activities in enterprises, natural persons conducting economic activity, as well as data on public roads and expenditures on public roads. The economic domain consists of 34 indicators, which represent almost 29% of the input data. Table 2 contains a list of all the indicators in the economic domain.

Data on the condition of the environment of the studied cities is presented in the third table. The environmental domain contains information on, among others: electricity consumption, carbon dioxide emissions and other air pollutants, forest cover and green belts, as well as municipal waste. The environmental domain comprises a total of 14 indicators, representing almost 12% of all input data; the full list of indicators is presented in Table 3.

The last institutional-political domain contains 12 indicators (i.e. only 10% of all input data) and describes the activity of the society in foundations, associations and social organizations, the structure of local legislative bodies, but also expenditures from municipal budgets and the effectiveness of local spatial planning. Table 4 contains an overview of all indicators in the institutional-political domain.

**Table 1.** List of indicators from social domain

No. of indicator	Thematic groups of indicators	Dimension [unit of measurement]
1	2	3
Demographic changes		
X1	Natural increase per 1000 population	total [-]
X2		deviation at points (percentage points) from the voivodeship value [-]
X3	Demographic dependency ratio	post-working age population per 100 persons of working age [person]
X4		post-working age population per 100 persons of working age – voivodeship=100 [%]
X5		non-working age population per 100 persons of working age [person]
X6		non-working age population per 100 persons of working age – voivodeship=100 [%]
X7		post-working age population per 100 persons of pre-working age [person]
X8		post-working age population per 100 persons of pre-working age – voivodeship=100 [%]

cont. Table 1

1	2	3
Public health		
X9	Infant deaths per 1000 live births	total [per mil]
X10		deviation at points from the voivodeship value [per mil]
X11	Deaths of people aged up to 65 years per 1000 population at this age	total [-]
X12		deviation at points from the voivodeship value [-]
Poverty and living conditions		
X13	People in households benefiting from the social assistance at domicile in the percentage of the total population	deviation from the voivodeship value [pp]
X14		total [%]
X15	Average monthly gross wages and salaries (economic entities which employ more than 9 persons)	total [PLN]
X16		voivodeship=100 [%]
X17	The average useful floor area of dwelling per capita	total [m <sup>2</sup> ]
X18		voivodeship=100 [%]
Education		
X19	Children covered by pre-school education in percentage of the total number of children at the age 3–5	total [%]
X20		total – deviation from the voivodeship value [pp]
X21		in rural areas – deviation from the voivodeship value [pp]
X22	Ratios the quality of education and the level of students' knowledge	Passing the exam maturity examination in the vocational upper secondary schools [%]
X23		Passing the exam maturity examination in the vocational upper secondary schools – deviation from the voivodeship value [pp]
X24		Passing the exam maturity examination in general secondary schools [%]
X25		Passing the exam maturity examination in the general secondary schools – deviation from the voivodeship value [pp]
Access to the labor market		
X26	Long-term unemployed persons in registered unemployed persons total	total [%]
X27		deviation from the voivodeship value [pp]
X28	Registered unemployed persons in relation to persons of working age	unemployed persons, females [%]
X29		unemployed persons, females - deviation from the voivodeship value [pp]
X30		unemployed persons with tertiary education, total [%]
X31		unemployed persons with tertiary education, total – deviation from the voivodeship value [pp]
X32	Graduates – registered unemployment (yet not working) in the percentage of the total registered unemployed persons	total [%]
X33		total – deviation from the voivodeship value [pp]
X34		graduates – females [%]
X35		graduates – females – deviation from the voivodeship value [pp]

cont. Table 1

1	2	3
X36	Registered unemployment rate	total [%]
X37		deviation from the voivodeship value [pp]
Sustainable consumption patterns		
X38	Number of passenger cars per 1000 population	total [pcs]
X39		voivodeship=100 [%]
X40	Consumption of water, electricity, and gas in households during the year per capita	electricity [kWh]
X41		electricity – voivodeship=100 [%]
X42		gas [m <sup>3</sup> ]
X43		gas – voivodeship=100 [%]
X44		water [m <sup>3</sup> ]
X45		water – voivodeship=100 [%]
Old-age income adequacy		
X46	Long-term unemployed persons aged 55–64 in relation to registered unemployed persons aged 55–64 total	total [%]
X47		deviation from the voivodeship value [pp]
X48	Persons at post-working age in households benefiting from social assistance at domicile in percentage of the total number of people at this age	total [%]
X49		deviation from the voivodeship value [pp]
Determinants of health		
X50	Persons injured in accidents at work per 1000 employed persons	total [person]
X51		deviation at points from the voivodeship value [person]
X52	Out-patient departments per 10 thousand population	total [facilities]
X53		voivodeship=100 [%]
Criminality		
X54	Rate of detectability of the delinquents of ascertained by Police crimes	total [%]
X55		deviation from the voivodeship value [pp]
X56	Ascertained by Police crimes total per 1000 population	total – voivodeship=100 [%]
X57		total [-]
Road accidents		
X58	Victims of road accidents per 100 thousand registered motor	injured [person]
X59		injured – voivodeship=100 [%]

Source: own elaboration, based on [Sustainable Development Indicators, 2017].

**Table 2.** List of indicators from economic domain

No. of indicator	Thematic groups of indicators	Dimension [unit of measurement]
1	2	3
Economic development		
X60	Investment outlays in enterprises (current prices; without economic entities employing up to 9 people) per capita at working age	per capita at working age [PLN]
X61		per capita at working age – voivodeship=100 [%]
X62	New-registered entities of the national economy recorded in the REGON register per 10 thousand population at the working age	total [-]
X63		voivodeship=100 [%]
X64	Expenditure on innovation activities in enterprises by a group of sections in the percentage of the total expenditure on innovation activities in enterprises (up to 9 employees)	agriculture, hunting and forestry; fishing – deviation from the voivodeship value [pp]
X65		industry and construction [%]
X66		industry and construction – deviation from the voivodeship value [pp]
X67		trade; repair of motor vehicles; transportation and storage; accommodation and catering; information and communication [%]
X68		trade; repair of motor vehicles; transportation and storage; accommodation and catering; information and communication – deviation from the voivodeship value [pp]
X69		financial and insurance activities; real estate activities [%]
X70		financial and insurance activities; real estate activities – deviation from the voivodeship value [pp]
X71		other services [%]
X72	other services – deviation from the voivodeship value [pp]	
Employment		
X73	Natural persons conducting economic activity per 100 persons of working age	total [-]
X74		voivodeship=100 [%]
X75	Entities by size classes per 10 thousand population at the working age	total [-]
X76		total – voivodeship=100 [%]
X77		Micro (up to 9 employees) [-]
X78		Micro (up to 9 employees) – voivodeship=100 [%]
X79		Small (from 10 to 49 employees) [-]
X80		Small (from 10 to 49 employees) – voivodeship=100 [%]
X81		Medium (from 50 to 249 employees) [-]
X82		Medium (from 50 to 249 employees) – voivodeship=100 [%]
X83	Large (over 250 employees) [-]	Large (over 250 employees) [-]
X84		Large (over 250 employees) – voivodeship=100 [%]
Transport		
X85	Length of bicycle lane	per 10 thousand km <sup>2</sup> [km]
X86		per 10 thousand km <sup>2</sup> – voivodeship=100 [%]
X87		per 10 thousand population [km]
X88		per 10 thousand population – voivodeship=100 [%]



cont. Table 2

1	2	3
X89	Length of local public roads per 100 km <sup>2</sup>	surfaced [km]
X90		surfaced – voivodeship=100 [%]
X91	Expenditures of municipalities on public roads in the percentage of their total expenditure	total [%]
X92		deviation from the voivodeship value [pp]
X93	Expenditures of districts on public roads in the percentage of their total expenditure	deviation from the voivodeship value [pp]

Source: own elaboration, based on [Sustainable Development Indicators, 2017].

**Table 3.** List of indicators from environmental domain

No. of indicator	Thematic groups of indicators	Dimension [unit of measurement]
Climate change		
X94	Emissions of carbon dioxide from plants especially noxious to air purity	total [t/y]
Energy		
X95	Electricity consumption per capita	total [kWh]
X96		total – voivodeship=100 [%]
X97		in urban areas [kWh]
X98		in urban areas – voivodeship=100 [%]
Air protection		
X99	Emissions of air pollutants from plants especially noxious to air purity	gases [t/y]
X100	Pollutants retained or neutralized in pollutant reduction systems in polluting plants especially noxious to air purity in percentage of the generated	particulates – deviation from the voivodeship value [pp]
Land use		
X101	Forest cover	total [%]
Biodiversity		
X102	Green belts in percentage of the total area	total [%]
X103		deviation from the voivodeship value [pp]
Waste management		
X104	Mixed municipal waste from household collected during the year per capita	total [kg]
X105		voivodeship=100 [%]
X106	Treated industrial and municipal wastewater in total	total [%]
X107	percentage of the total value of the industrial and municipal wastewater requiring treatment	deviation from the voivodeship value [pp]

Source: own elaboration, based on [Sustainable Development Indicators, 2017].

**Table 4.** List of indicators from institutional–political domain

No. of indicator	Thematic groups of indicators	Dimension [unit of measurement]
Openness and participation		
X108	Foundations, associations and social organizations per 10 thousand population	total [-]
X109		voivodeship=100 [%]
X110	Structure of councilors in the organs (legislative bodies) in municipalities and districts	females [%]
X111		females – deviation from the voivodeship value [pp]
X112		people with higher education [%]
X113		people with higher education – deviation from the voivodeship value [pp]
Economic instruments		
X114	Expenditure from the budgets of municipalities and districts on public debt in 1000 PLN per total revenue budgets of municipalities and districts	total, (include municipalities and powiats) [PLN]
X115		total, (include municipalities and powiats) – voivodeship=100 [%]
X116	Investment expenditures of municipalities and districts in percentage of their total expenditure	total, (include municipalities and powiats) [%]
X117		total, (include municipalities and powiats) – deviation from the voivodeship value [pp]
X118	Area covered by the local spatial development plans in percentage of the total area	total [%]
X119		deviation from the voivodeship value [pp]

Source: own elaboration, based on [Sustainable Development Indicators, 2017].

For the collected diagnostic variables, it should be examined whether these variables are characterized by sufficiently high variability by eliminating quasi-constant variables. For this purpose, the coefficient of variation  $V$  can be calculated for each  $j$ -th variable. Its value is a relative measure of dispersion, and it is calculated by using the Equation (1) below.

$$V_j = \frac{S_j}{\bar{x}_j}, (j = 1, \dots, m), \quad (1)$$

where:  $\bar{x}_j$  – the arithmetic mean of the  $j$ -th variable (2),  $S_j$  – standard deviation for the  $j$ -th variable (2).

$$\bar{x}_j = n^{-1} \sum_{i=1}^n x_{ij}, (i = 1, \dots, n);$$

$$S_j = \sqrt{n^{-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}, \quad (2)$$

From the set of variables, unequal variables can be eliminated.

$$|V_j| \leq V^*, \quad (3)$$

where  $V^*$  is the critical value of the variation coefficient. The value of  $V^*$  was arbitrarily set at 0.10.

Afterward, the strength of the relationship between the other variables should be tested. For this purpose, the correlation between variables must be determined with the value of the Pearson coefficient. Highly correlated variables are removed from the data set (Pearson's coefficient  $>0.7$ ) [Schober et al., 2018].

Covariance is a measure of the joint variability of two random variables. The covariance of variables shows how variables are linearly related to each other. Positive covariance indicates a positive linear relationship between variables, while negative covariance indicates the opposite. If the variables are not linearly related, the covariance value is close to zero. The covariances must be computed for the analyzed variables.

## Determination of the sustainable development level using the proposed method

Each variable has been separately classified using the Jenks optimization algorithm. Separate classification of each of the variables will ensure the reliability of the method – the range of each variable may vary significantly. Using Natural Breaks, a class has been assigned to each city and the values of these classes have been averaged for each city to obtain an indicator of Sustainable Development Level (SDL).

The method proposed above was used to compute the Sustainable Development Level (SDL) for all 66 district cities in Poland. These are the largest cities in the country. Its locations are shown on Figure 1.

## Determination of the sustainable development level using Classic Ranking and Hellwig's methods

Analogous research was done by using the Classic Ranking method and Hellwig's method, then the results were examined using the Pearson correlation coefficient. Ranking of Polish cities in terms of the level of sustainable development was created. This approach is used e.g. on Eurostat's SDI websites [EUROSTAT, 2019]. The Classic Ranking was computed for the same variables for which the sustainable development level was computed using the proposed method of classifying objects based on Natural Breaks; for the value of each indicator in each of the analyzed cities a number was assigned based on the position. This means that position 1 is the highest and position 66 the lowest. In the case of indicators described as stimulants, the city with the highest value of the indicator was given the number 1, the city with the lowest – 66 (the situation is the opposite for indicators described as destimulants). The results for the cities were averaged. The compilation of the obtained results and their interpretation sought the basis for discussion and conclusions in the last chapter.

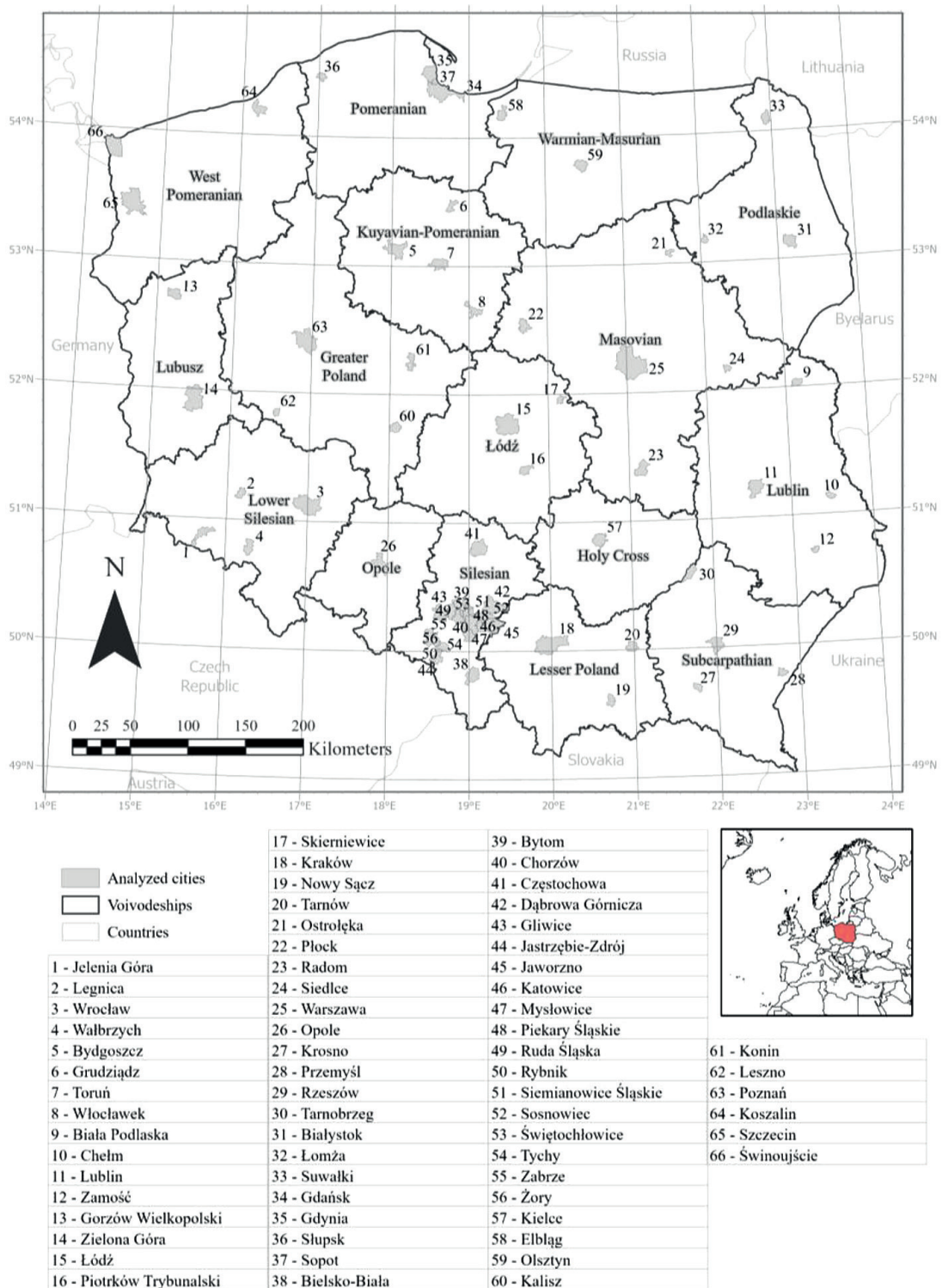
Taxonomic methods are most often used to determine the level of development of a given area or objects. These are statistical methods used to classify objects described by many of its properties. Analysis of the literature indicated that, in the case of research similar to the research carried out in this article, linear ordering methods are most often used. As a result, it was decided to choose the "Hellwig's method", a method proposed in 1968 by Polish scientist Zdzisław Hellwig. This method is commonly used in this type of research [Dorożyński et al., 2019, Jaworska & Luty, 2009, Łogwiniuk, 2011, Malina, 2020, Niemczyk, 2001, Podstawka & Suchodolski, 2018, Pomianek, 2010, Rzasa et al., 2019, Salamon, 2005, Senetra & Szarek-Iwaniuk, 2020, Sołek & Sowa, 2019, Stec, 2012, Ziemiańczyk, 2010]. Hellwig's method is based on the computation of a synthetic development index which allows for presenting the situation of diversity in the level of the studied phenomenon, covering many categories: economic, social, ecological, and spatial [Hellwig, 1968, Nowak, 1990]. The adopted methodology of the research procedure is characterized by great transparency, as it makes it possible to present the results with a single numerical value. This is a great advantage of this method and a premise for its selection [Ilnicki, 2002]. The construction of a synthetic developmental index requires several stages, starting from the selection of a set of objects and diagnostic variables, through normalization of features, determination of stimulants and destimulants, to the computation of the index value, as a distance from the constructed developmental index.

The numerical description of the set of objects can be presented in the form of an observation matrix  $\mathbf{X}$ , taking the form of Equation (4) below.

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}, \quad (4)$$

where  $x_{ij}$  means the value of the  $j$ -th variable for the  $i$ -th object ( $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ ).

In the first step, after selecting the variables, according to the rules described in "Selection



**Fig. 1.** Location of the studied cities in relation to Polish voivodeships  
 Source: own elaboration using ArcGIS Pro 2.5, based on [Polish National Register of Borders, 2020].

of diagnostic variables” section, the variables must be unified. To unify variables, the characteristics should be normalized by standardizing it, according to Equation (5).

$$Z_{ij} = \frac{(x_{ij} - \bar{x}_j)}{s_j}, (j = 1, \dots, m), \quad (5)$$

where:  $\bar{x}_j$  is the arithmetic mean of  $j$ -th variable and  $s_j$  is the standard deviation for the  $j$ -th variable. This way, a matrix of standard values of the  $Z$  characteristics is obtained in Equation (6) below.

$$Z = \begin{bmatrix} z_{11} & \dots & z_{1m} \\ \vdots & \ddots & \vdots \\ z_{n1} & \dots & z_{nm} \end{bmatrix}, \quad (6)$$

where  $z_{ij}$  is a standardized value of  $x_{ij}$ .

The matrix (6) formed is the basis for determining the reference object  $P_0$ . It is an abstract object (eg. city) with standardized values  $z_{01}, z_{02}, \dots, z_{0j}$ , where:

$$\left( \begin{array}{l} z_{0j} = \max_i z_{ij}, \text{ when } X_j \text{ is a stimulant} \\ z_{0j} = \min_i z_{ij}, \text{ when } X_j \text{ is a destimulant} \end{array} \right), \quad (7)$$

The  $P_0$  object obtained in this way is treated as a development pattern.

In the next step, the Euclidean distances of the tested objects from the determined pattern should be calculated. This can be completed based on Equation (8).

$$D_{i0} = \sqrt{\sum_{j=1}^m (z_{ij} - z_{0j})^2}, \quad (8)$$

For the  $D_{10}, D_{20}, \dots, D_{n0}$  distance values obtained in this way, the average value should be calculated (9).

$$\bar{D}_0 = n^{-1} \sum_{i=1}^n D_{i0}, \quad (9)$$

As well as standard deviation (10):

$$S_0 = \sqrt{n^{-1} \sum_{i=1}^n (D_{i0} - \bar{D}_0)^2}, \quad (10)$$

The level of sustainable development is obtained from Equation (11) below.

$$d_i = 1 - \frac{D_{i0}}{D_0}, \quad (11)$$

where:

$$D_0 = \bar{D}_0 + 2S_0, \quad (12)$$

A string of  $d_1, d_2, \dots, d_n$  values is obtained in this way, using the range [0.1].

The higher the measure of the  $d_i$  value of the tested object (i.e., its values are close to the pattern), the higher its level of development is. The lower the  $d_i$  value is (i.e., the values of the tested object are further away from the pattern), the lower its level of development is.

Two parameters of the taxonomic measure can be used to classify the examined objects, according to the level of sustainable development: geometric mean ( $\bar{d}_i$ ) and standard deviation ( $S_{di}$ ). Six sustainable development classes of cities can be distinguished in this way, depending on the value of  $d_i$ :

1. Sixth class (the lowest level of development):  
 $d_i < \bar{d}_i - 2S_{di}$
2. Fifth class (low level of development):  
 $\bar{d}_i - 2S_{di} \leq d_i < \bar{d}_i - S_{di}$
3. Fourth class (medium level of development):  
 $\bar{d}_i - S_{di} \leq d_i < \bar{d}_i$
4. Third class (medium–high level of development):  
 $\bar{d}_i \leq d_i < \bar{d}_i + S_{di}$
5. Second class (high level of development):  
 $\bar{d}_i + S_{di} \leq d_i < \bar{d}_i + 2S_{di}$
6. First class (the highest level of development):  
 $d_i \geq \bar{d}_i + 2S_{di}$

Determination of classes is usually a necessary element for analyses using Hellwig’s method, but in the case described in the article it would be an unnecessary rounding – instead, the exact  $d_i$  values for each city will be used for the analysis using the Pearson correlation coefficient.

## RESULTS

### Preparation of the data

The process of verifying the indicators for usefulness and usability in the proposed method, described in “Selection of diagnostic variables”

section, involves the rejection of indicators based on the requirements of the method and consists of four steps:

1. Pre-selection of indicators based on dimensions and the possibility of using it for national analysis (60 rejected indicators).
2. Rejection of indicators with low variation coefficient (5 rejected indicators).
3. Rejection of indicators with a high level of correlation – Pearson’s linear correlation analysis (17 rejected indicators).
4. Verification of linearity of diagnostic variables – covariance analysis (2 rejected indicators).

All indicators for analysis have been marked continuously from  $X_1$  to  $X_{119}$  – the numbering will be preserved until the end of the study to clearly depict the rejected indicators in each of the four steps. In the first step, the usage of indicators was examined on the basis of thematic groups and dimensions. The only reason for the rejection of the indicators turned out to be a relation to data for the voivodeship – all 60 indicators were rejected on this basis. Originally, the data is published in the form of data packages for a single voivodeship, in which each district is one row. Referring data for districts to the average value for a voivodeship is relevant in the case of analyzing a voivodeship; in the case of synthetic analysis of individual districts from different voivodeships in the whole country, the use of such data has no substantive value. The rejected indicators belonged to all domains.

The next step of rejection of indicators was the analysis of the variation coefficient. Five indicators for which the computed value of this coefficient was in the range  $\langle -0.1; 0.1 \rangle$  were rejected. All rejected indicators belonged to the social domain.

The implementation of Pearson’s linear correlation analysis was the third step of verifying the indicators. The generated matrix of coefficients was used to reject 17 indicators whose correlation coefficients were higher than 0.7. The rejected indicators belonged to all domains.

The last step of the rejection of indicators was a covariance analysis, which resulted in the rejection of two indicators from the social domain. A matrix

of covariance coefficients was generated and the basis for the rejection was the result in the range  $\langle -0.1; 0.1 \rangle$ .

In the process of verifying the input data, a total of 84 indicators out of 119 were rejected. In the first step – the pre-selection of indicators based on dimensions – the following indicators were rejected:  $X_2, X_4, X_6, X_8, X_{10}, X_{12}, X_{13}, X_{16}, X_{18}, X_{20}, X_{21}, X_{23}, X_{25}, X_{27}, X_{29}, X_{31}, X_{33}, X_{35}, X_{37}, X_{39}, X_{41}, X_{43}, X_{45}, X_{47}, X_{49}, X_{51}, X_{53}, X_{55}, X_{56}, X_{59}, X_{61}, X_{63}, X_{64}, X_{66}, X_{68}, X_{70}, X_{72}, X_{74}, X_{76}, X_{78}, X_{80}, X_{82}, X_{84}, X_{86}, X_{88}, X_{90}, X_{92}, X_{93}, X_{96}, X_{98}, X_{100}, X_{103}, X_{105}, X_{107}, X_{109}, X_{111}, X_{113}, X_{115}, X_{117}, X_{119}$ . Analysis of the variation coefficient led to the rejection of the indicators:  $X_5,$

**Table 5.** Overview of rejected indicators

Area/domain	Step 1	Step 2	Step 3	Step 4
Demographic changes	4	1	1	1
Public health	2	0	1	1
Poverty and living conditions	3	1	0	0
Education	4	3	0	0
Access to labor market	6	0	4	0
Sustainable consumption patterns	4	0	0	0
Old-age income adequacy	2	0	1	0
Determinants of health	2	0	0	0
Criminality	2	0	0	0
Road accidents	1	0	0	0
<b>Social domain</b>	<b>30</b>	<b>5</b>	<b>7</b>	<b>2</b>
Economic development	7	0	2	0
Employment	6	0	6	0
Transport	5	0	0	0
<b>Economic domain</b>	<b>18</b>	<b>0</b>	<b>8</b>	<b>0</b>
Climate change	0	0	0	0
Energy	2	0	0	0
Air protection	1	0	1	0
Land use	0	0	0	0
Biodiversity	1	0	0	0
Waste management	2	0	0	0
<b>Environmental domain</b>	<b>6</b>	<b>0</b>	<b>1</b>	<b>0</b>
Openness and participation	3	0	1	0
Economic instruments	3	0	0	0
<b>Institutional-political domain</b>	<b>6</b>	<b>0</b>	<b>1</b>	<b>0</b>
Overall	60	5	17	2

Source: own elaboration.

$X_{17}, X_{19}, X_{22}, X_{24}$ . On the basis of Pearson’s linear correlation analysis, the following indicators were rejected:  $X_7, X_{11}, X_{28}, X_{30}, X_{34}, X_{36}, X_{46}, X_{62}, X_{67}, X_{73}, X_{75}, X_{77}, X_{79}, X_{81}, X_{83}, X_{99}, X_{108}$ . The last step, the analysis of covariance has resulted in the rejection of indicators  $X_1$  and  $X_9$ . The overview of rejected indicators by steps and areas is presented in Table 5.

The last step of data preparation is to assign an explanatory variable to each variable. A variable can be a stimulant or a destimulant; for a stimulant, an increase in the value of the explanatory variable leads to an increase in the variable, for a destimulant, an increase in the value of the explanatory variable leads to a decrease in the variable.

### Results for the proposed method

Step “Preparation of the data” was performed for 119 indicators characterizing sustainable development of Poland’s biggest cities. After appropriate rejection of the variables, the remaining variables became the basis for classification using the proposed method. The classification process described in “Determination of the sustainable development level using the proposed method” section has become the basis for determining Sustainable Development Level (SDL) for the analyzed cities. The results for cities with the highest and the lowest SDL values are presented in Table 6. All results are shown in Table 9 below.

**Table 6.** Cities with the highest and the lowest SDL values

TERYT code	City name	SDL using Natural Breaks
1261000	Kraków	4.2857
3064000	Poznań	4.1429
0264000	Wrocław	4.0857
2261000	Gdańsk	4.0857
1465000	Warszawa	4.0286
...		
0462000	Grudziądz	3.0286
0464000	Włocławek	3.0286
2478000	Zabrze	3.0286
1461000	Ostrołęka	2.9714
2063000	Suwałki	2.9143

Source: own elaboration.

### Results for Classic Ranking and Hellwig’s methods

The process described in “Determination of the sustainable development level using Classic Ranking and Hellwig’s methods” section allowed to compute Sustainable Development Level (SDL) using Hellwig’s taxonomic method, Table 7 shows the results of the research – results for cities with the highest and lowest  $d_i$  value. All results are shown in Table 9 below.

In tables above, the best results section contains four identical cities, while the worst results section contains two identical cities. The last step was

**Table 7.** Cities with the highest and the lowest  $d_i$  values

TERYT code	City name	SDL using Hellwig’s method ( $d_i$ value)
3064000	Poznań	0.3107
1465000	Warszawa	0.3093
2264000	Sopot	0.3020
0264000	Wrocław	0.3007
1261000	Kraków	0.2754
...		
0464000	Włocławek	0.0477
2063000	Suwałki	0.0471
1062000	Piotrków Trybunalski	0.0465
1262000	Nowy Sącz	0.0338
1864000	Tarnobrzeg	0.0264

Source: own elaboration.

**Table 8.** Cities with the highest and the lowest SDL values, computed using Classic Ranking

TERYT code	City name	SDL using Classic Ranking
0264000	Wrocław	42.5143
1261000	Kraków	42.4000
3064000	Poznań	42.2286
2261000	Gdańsk	40.6571
2264000	Sopot	39.9714
...		
0664000	Zamość	27.1143
0462000	Grudziądz	26.3714
0464000	Włocławek	26.0857
2063000	Suwałki	25.3714
1461000	Ostrołęka	24.9714

Source: own elaboration.

to compute the SDL using the Classic Ranking. The following Table 8 contains the results for the cities with the highest and the lowest score. All results are shown in Table 9 below.

In the case of results from the Classic Ranking, all five cities in the top of the table are identical, and for the bottom of the table – four.

## Final results

The results of the research – SDL using the proposed method, Hellwig’s method, and Classic Ranking method – are summarized in Table 9.

The highest SDL value was obtained by the city of Kraków, definitely ahead of other cities such as: Poznań, Wrocław, Gdańsk, Warszawa, and Sopot. Both the highest and the lowest results are described by a lack of concentration in one voivodeship – five cities with the highest SDL values are located in different voivodeships, as well as four cities with the lowest values. Table 10 below contains the computed basic descriptive statistics indicators for the results of SDL using different methods.

In order to compare the results obtained with the proposed method with the results obtained with the Hellwig’s method and the Classic Ranking method,

**Table 9.** Results of the research – Sustainable Development Level of analyzed cities

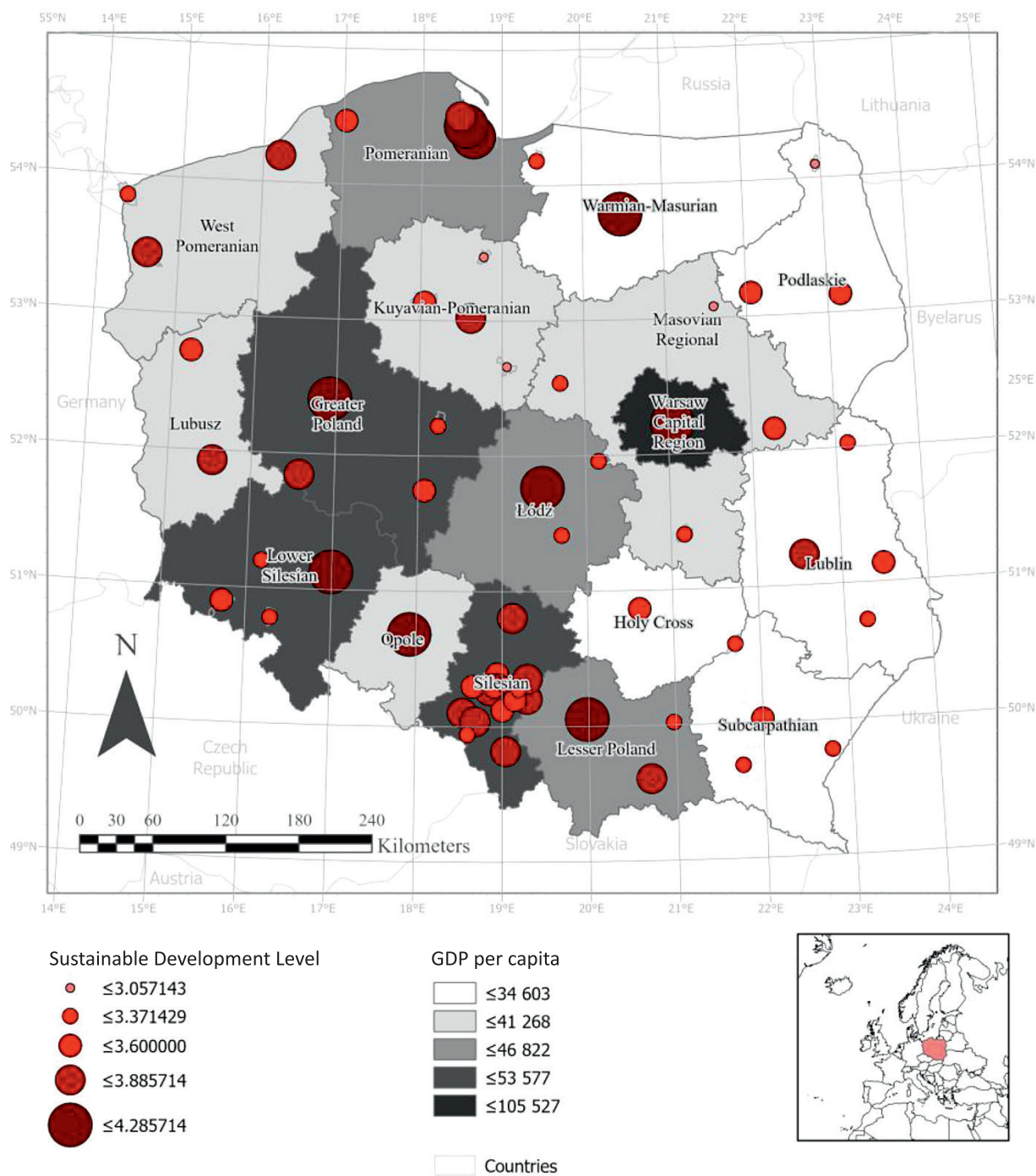
TERYT code	City	SDL using the proposed method		SDL using Hellwig’s method (di value)		SDL using Classic Ranking	
		value	position	value	position	value	position
1	2	3	4	5	6	7	8
1261000	Kraków	4.29	1	0.28	5	42.40	2
3064000	Poznań	4.14	2	0.31	1	42.23	3
0264000	Wrocław	4.09	3	0.30	4	42.51	1
2261000	Gdańsk	4.09	3	0.27	6	40.66	4
1465000	Warszawa	4.03	5	0.31	2	39.91	6
2264000	Sopot	4.03	5	0.30	3	39.97	5
1661000	Opole	4.00	7	0.21	8	39.63	7
1061000	Łódź	3.97	8	0.18	17	37.97	9
2862000	Olsztyn	3.94	9	0.17	21	38.86	8
2262000	Gdynia	3.89	10	0.18	19	36.17	15
2464000	Częstochowa	3.86	11	0.16	22	37.46	10
2473000	Rybnik	3.86	11	0.11	41	37.14	11
3063000	Leszno	3.80	13	0.18	20	35.89	17
2479000	Żory	3.77	14	0.14	31	35.51	20
3262000	Szczecin	3.74	15	0.18	16	35.94	16
1262000	Nowy Sącz	3.71	16	0.03	65	33.14	28
2469000	Katowice	3.71	16	0.24	7	36.43	12
0463000	Toruń	3.69	18	0.21	9	35.69	18
0862000	Zielona Góra	3.69	18	0.14	30	35.11	22
2461000	Bielsko-Biała	3.69	18	0.19	14	35.49	21
2465000	Dąbrowa Górnicza	3.69	18	0.14	32	35.63	19
2468000	Jaworzno	3.69	18	0.19	13	36.20	14
2472000	Ruda Śląska	3.69	18	0.15	29	36.40	13
0663000	Lublin	3.66	24	0.20	12	34.91	23
3261000	Koszalin	3.66	24	0.15	26	33.91	26



cont. Table 9

1	2	3	4	5	6	7	8
2061000	Białystok	3.60	26	0.14	33	32.37	34
2471000	Piekary Śląskie	3.60	26	0.11	42	32.26	36
2477000	Tychy	3.60	26	0.21	10	34.74	24
0662000	Chełm	3.57	29	0.07	52	32.80	32
1863000	Rzeszów	3.57	29	0.16	25	33.06	29
2463000	Chorzów	3.57	29	0.19	15	33.83	27
0861000	Gorzów Wielkop.	3.51	32	0.16	24	34.20	25
2263000	Słupsk	3.51	32	0.13	34	33.03	30
2470000	Mysłowice	3.51	32	0.10	45	32.57	33
0261000	Jelenia Góra	3.49	35	0.15	27	32.91	31
2062000	Łomża	3.49	35	0.11	39	31.37	40
1464000	Siedlce	3.46	37	0.18	18	31.89	38
3061000	Kalisz	3.46	37	0.16	23	31.31	42
2661000	Kielce	3.43	39	0.10	46	32.17	37
0461000	Bydgoszcz	3.40	40	0.15	28	30.69	44
2466000	Gliwice	3.40	40	0.21	11	32.31	35
2476000	Świętochłowice	3.40	40	0.08	49	29.83	49
1862000	Przemyśl	3.37	43	0.10	47	31.54	39
2861000	Elbląg	3.37	43	0.09	48	30.37	46
0262000	Legnica	3.34	45	0.13	36	31.06	43
1063000	Skierniewice	3.34	45	0.13	35	31.34	41
1463000	Radom	3.34	45	0.10	44	30.06	48
3062000	Konin	3.34	45	0.07	57	29.71	50
0661000	Biała Podlaska	3.31	49	0.07	53	30.51	45
0265000	Wałbrzych	3.26	50	0.08	50	30.29	47
1864000	Tarnobrzeg	3.26	50	0.03	66	28.66	53
1062000	Piotrków Tryb.	3.23	52	0.05	64	28.66	53
1462000	Płock	3.23	52	0.12	38	28.83	52
2467000	Jastrzębie-Zdrój	3.23	52	0.08	51	27.97	56
2474000	Siemianowice Śl.	3.23	52	0.07	58	28.60	55
3263000	Świnoujście	3.23	52	0.13	37	27.23	61
1861000	Krosno	3.20	57	0.10	43	27.57	58
0664000	Zamość	3.17	58	0.11	40	27.11	62
1263000	Tarnów	3.17	58	0.07	56	27.83	57
2475000	Sosnowiec	3.14	60	0.06	60	29.40	51
2462000	Bytom	3.06	61	0.07	55	27.43	60
0462000	Grudziądz	3.03	62	0.07	59	26.37	63
0464000	Włocławek	3.03	62	0.05	62	26.09	64
2478000	Zabrze	3.03	62	0.07	54	27.54	59
1461000	Ostrołęka	2.97	65	0.05	61	24.97	66
2063000	Suwałki	2.91	66	0.05	63	25.37	65

Source: own elaboration.



**Fig. 2.** Sustainable Development Level and GDP per capita  
 Source: own elaboration using ArcGIS Pro 2.5, based on results and [Urząd Statystyczny w Katowicach, 2018].

**Table 10.** Descriptive statistics for SDL values

Descriptive statistics	SDL using the proposed method	SDL using Hellwig's method (di value)	SDL using Classic Ranking
Mean	3.5260	0.1404	32.8944
Standard error	0.0390	0.0087	0.5387
Median	3.5143	0.1340	32.4714
Standard deviation	0.3170	0.0707	4.3763
Variance	0.1005	0.0050	19.1519
Kurtosis	-0.5186	0.0382	-0.5006
Skewing	0.2405	0.6917	0.3328
Minimum	2.9143	0.0264	24.9714
Maximum	4.2857	0.3107	42.5143
Range	1.3714	0.2843	17.5429

Source: own elaboration.

**Table 11.** Pearson correlation coefficient for SDL values

	Proposed method	Hellwig's Method	Classic Ranking Method
Proposed Method		0.81	0.98
Hellwig's Method			0.85
Classic Ranking Method			

Source: own elaboration.

the correlation between the results was examined, using the Pearson correlation coefficient. Table 11 contains a set of correlation coefficients.

The results show a high correlation; a greater correlation between the proposed method and the Classic Ranking method can be observed – the value of Pearson's correlation coefficient is as high as approximately 0.98.

Sustainable development level values were used to develop a spatial distribution map using ArcGIS Pro 2.5 software (Fig. 2). Geographic Information System was used to process and visualize data – this is a powerful tool to support decision-making processes in many fields of science [Ciski & Rzasa, 2018]. The values of SDL were visualized using the graduated symbols method. For further discussion of the results, as a background by using the cartogram method the GDP per capita by NUTS units was displayed (the NUTS division is identical to the division into voivodeships, with the exception of the Masovian

Voivodeship divided into the Masovian Regional and the Warsaw Capital Region).

## CONCLUSIONS

The research conducted in this article and its results confirms the possibility of applying the proposed method of classifying objects based on natural breaks in analyses of the level of sustainable development of cities. The ranking of 66 Polish cities in terms of its sustainable development level was obtained. Therefore, the assumed aim of the research has been achieved.

The authors decided to verify the correlation of the results obtained with those from Hellwig's Method and the Classic Ranking. A higher correlation was observed with the Classical Ranking (Pearson's correlation coefficient reached approximately 0.98), which indicates a very strong correlation between these methods. The proposed method and Hellwig's method indicate lower but still high correlation (Pearson's correlation coefficient was approximately 0.81). By analyzing Hellwig's method, it can be indicated that it takes greater account of the variation in values within the variables for different objects. The classic ranking refers only to the positions on the list. The Hellwig's method and the Classic Ranking method are commonly used in this type of analysis and research; the fact of receiving highly correlated results confirms the validity and reliability of the proposed method. An overall trend of the results is confirmed by such a high result (e.g. the same four cities are in the top five). However, as can be seen in Table 9 (columns 5 and 7), the positions in the ranking of individual cities differ quite clearly from each other.

The results obtained by the Authors are also confirmed by other authors' research, using other methods (other taxonomic methods, multidimensional analysis, etc.) [Adamowicz & Janulewicz, 2016, Koszel & Bartkowiak, 2018, Laskowska & Dańska-Borsiak, 2018, Łuczak & Kurzawa, 2017, Matras, 2017, Mięka, 2020]. The very order of the cities in the ranking is obviously different using different methods, but some general trend is maintained.

The number of indicators rejected in the research may raise some doubts (the final analysis took into account 35 out of 119). However, this is due to the requirements for the proper selection of variables. In this regard, the authors have adjusted to the requirements related to the statistical elaboration in order to obtain reliable and credible results. A large number of original indicators were more related to the form of its publication in different dimensions than to thematic diversity. The 35 indicators adopted in the research were fully representative of the level of sustainable development and covered all the domains analyzed thematically. Even the percentage share of indicators from individual domains (taken into account in the computation in relation to the original number) was maintained with great approximation, which also confirms the validity of the selection of the thematic coverage by variables of sustainable development concepts.

The results obtained were certainly influenced by the economic condition of the voivodeships (expressed as GDP per capita) in which the analyzed cities are located. Analyzing the cartodiagram presented in Figure 2, it can be seen that the four out of five cities that obtained the highest level of sustainable development in the research are located in the voivodeships with the two highest classes of GDP/inhabitant. This is also confirmed by the results obtained by other authors in the analysis of sustainable development, as well as socio-economic development performed for the voivodeships [Drabarczyk, 2017, Klóska, 2017, Michoń, 2017, Misiewicz et al., 2019]. The obtained results were also influenced by the density of cities in particular voivodeships. In Silesian Voivodeship there are 19 cities, and in Opole and Holy Cross Voivodeships – only one each.

One can observe a certain correlation between the results obtained and the size of cities expressed in terms of their population. However, this relationship is not a linear relationship indicating that larger cities are higher in the ranking. There are several deviations from this claim. Among the 66 analyzed Polish cities, one can distinguish: 5 cities with a population of up to 50,000 inhabitants, 22 – with

a population of 50,000–100,000 inhabitants, 28 – with a population of 100,000–250,000, 6 – with a population of 250,000–500,000, 4 – with a population of 500,000–1 million and one – with a population over 1 million. Among the five cities which obtained the highest sustainable development level are: Kraków, Poznań and Wrocław (500,000–1 million residents), Gdańsk (250,000–500,000 residents), and Warsaw (over 1 million residents). Among the five cities with the lowest sustainable development level are: Grudziądz, Ostrołęka, and Suwałki (50,000–100,000 residents), and Włocławek and Zabrze (100,000–250,000 residents).

The methodology proposed by the authors may be very widely used in various studies. It can be used to compare not only cities but also e.g. other territorial division units both in a given country, or to compare different countries. The thematic areas of the analyses may also be different. The proposed methodology can also be used e.g. to assess the differences in the level of economic and social development, or the smart-city level – very popular in recent years.

The results of the analyses obtained by the Authors, as well as the research methodology itself, can be used to create various rankings of cities or local government units. Such results can be e.g. used by local authorities in activities related to the promotion of a given city, attracting new residents or investors, or other activities related to territorial marketing. On the other hand, for cities with the worst results, it may constitute a developmental stimulus, justifying e.g. taking various not necessarily popular decisions by local authorities.

**Author contributions:** authors have given approval to the final version of the article. Authors contributed to this work as follows: K.R. and developed the concept and designed the study, M.C. collected the data, K.R. and M.C. analysed and interpreted the data, K.R. and M.C. prepared draft of article, K.R. and M.C. revised the article critically for important intellectual content.

**Funding:** This research received no external funding.

## REFERENCES

- Act of 27 April 2001 Environmental Protection Law, 62 Polish Journal of Law § (2001) (Poland).
- Adamowicz, M., Janulewicz, P. (2016). Ocena poziomu rozwoju dużych miast w Polsce – Wyniki badań wykorzystujących wzorzec rozwoju Hellwiga [Evaluation of the level of development of large cities in Poland – results of studies using the Hellwig's development model]. *Studia Regionalne i Lokalne [Regional and Local Studies]*, 64(2), pp. 68–86. doi:10.7366/1509499526404.
- Aggarwal, C.C. (2014). *Data classification: Algorithms and Applications*. London, Chapman and Hall/CRC.
- Alkhalidi, A., Qoaidar, L., Khashman, A., Al-Alami, A.R., Jiryas, S. (2018). Energy and Water as Indicators for Sustainable City Site Selection and Design in Jordan using smart grid. *Sustainable Cities and Society*, 37, pp. 125–132. doi:10.1016/j.scs.2017.10.037.
- Aly, M. (2005). *Survey on multiclass classification methods*. Neural Networks, Caltech.
- Antonopoulos, E. (2018). Sustainable Urban Development in Greece Using Integrated Territorial Investments. *European Structural & Investment Funds Journal*, 6(1), pp. 28–34.
- Atkinson, G., Dubourg, R., Hamilton, K., Munasinghe, M., Pearce, D., Young, C. (1997). *Measuring Sustainable Development: Macroeconomics and the Environment*. Cheltenham, Edward Elgar Publishing.
- Azapagic, A., Perdan, S. (2000). Indicators of sustainable development for industry: A general framework. *Process Safety and Environmental Protection*, 78(4). doi:10.1205/095758200530763.
- Bąk, A. (2016). Porządkowanie liniowe obiektów metodą Hellwiga i TOPSIS – analiza porównawcza [Linear ordering of objects using Hellwig and TOPSIS methods – a comparative analysis]. *Taksonomia 26 Klasyfikacja i analiza danych – teoria i zastosowania [Taxonomy 26 Classification and Data Analysis – Theory and Applications]*, pp. 22–31. doi:10.15611/pn.2016.426.02.
- Bąk, I., Cheba, K. (2018). Syntetyczna ocena dysproporcji zrównoważonego rozwoju krajów Unii Europejskiej [Synthetic evaluation of disproportions of sustainable development of the European Union countries]. *Wiadomości Statystyczne [Statistical News]*, 5(684), pp. 37–55.
- Bell, S., Morse, S. (2018). *Routledge Handbook of Sustainability Indicators*. Londyn, Routledge.
- Bossel, H. (1999). *Indicators for sustainable development: theory, method, applications*. International Institute for Sustainable Development, Winnipeg.
- Chen, J., Yang, S., Li, H., Zhang, B., Lv, J. (2013). Research on geographical environment unit division based on the method of natural breaks (Jenks). *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives*, 40(4W3), pp. 47–50. doi:10.5194/isprsarchives-XL-4-W3-47-2013.
- Ciski, M., Rzasa, K. (2018). Cultural Parks in the Spatial Planning System in Poland. *Proceedings – 2018 Baltic Geodetic Congress, BGC-Geomatics 2018*, pp. 145–149. doi:10.1109/BGC-Geomatics.2018.00033.
- Ciski, M., Rzasa, K., Ogryzek, M. (2019). Use of GIS tools in sustainable heritage management-the importance of data generalization in spatial modeling. *Sustainability*, 11(20). doi:10.3390/su11205616.
- Czerwińska, M. (2002). Zastosowanie metod taksonomicznych w klasyfikacji krajów Unii Europejskiej z punktu widzenia poziomu ich rozwoju gospodarczego [Application of taxonomic methods in the classification of European Union countries from the point of view of their level of economic development]. *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie [Scientific Journals of the Cracow University of Economics]*, 575(284), pp. 149–161.
- Dasgupta, P. (2007). Measuring sustainable development: Theory and application. *Asian Development Review*, 24(1), pp. 1–10.
- Dorożyński, T., Dobrowolska, B., Kuna-Marszałek, A. (2019). Institutional Quality As A Determinant of FDI Inflow: The case of Central and Eastern European Countries. *Journal of Management and Financial Sciences*, XII(36), pp. 103–122.
- Drabarczyk, K. (2017). Zrównoważony rozwój województw – analiza porównawcza [Sustainable Development of Regions – A Comparative Analysis]. *Zeszyty Naukowe Politechniki Częstochowskiej Zarządzanie*, 25(2), pp. 23–34. doi:10.17512/znpcz.2017.1.2.02.
- Drastichová, M. (2017). *Measuring Sustainability of the EU Countries with the Selected Indices*. The International Scientific Conference INPROFORUM 2017, 11–20. České Budějovice.
- European Commission. (2015). Indicators for sustainable cities. *Science for Environment Policy*. doi:10.2779/61700.
- European Commission. (2016). *Communication from the Commission to the European Parliament, the Council,*

- the European Economic and Social Committee and the Committee of the Regions*. Strasbourg.
- EUROSTAT. (2019). Indicators – Eurostat. <https://ec.europa.eu/eurostat/web/sdi/indicators>, date: 15.07.2020.
- Fierich, J. (1957). Próba zastosowania metod taksonomicznych do rejonizacji systemów rolniczych w województwie krakowskim [An attempt to use taxonomic methods for regionalization of agricultural systems in the Krakowskie Voivodeship]. *Mysł Gospodarcza [Economic Thought]*, 1.
- Gorzelał, G. (1979). Dobór zmiennych w statystycznej analizie porównawczej – metody (I) [Selection of variables in comparative statistical analysis – methods (I)]. *Wiadomości Statystyczne [Statistical News]*, 3, pp. 17–21.
- Gorzelał, G. (1981). Statystyczna analiza porównawcza – teoria a praktyka [Statistical comparative analysis – theory versus practice]. *Wiadomości Statystyczne [Statistical News]*, 8.
- Grabiński, T., Wydymus, S., Zeliaś, A. (1989). *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych [Methods of numerical taxonomy in modeling of socio-economic phenomena]*. Warszawa, PWN.
- Hák, T., Janoušková, S., Moldan, B. (2016). Sustainable Development Goals: A need for relevant indicators. *Ecological Indicators*, 60. doi:10.1016/j.ecolind.2015.08.003.
- Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom rozwoju i strukturę kwalifikowanych kadr [Application of the taxonomic method to the typological division of countries by the level of development and the structure of qualified human resources]. *Przegląd Statystyczny [Statistical Review]*, 4, pp. 307–327.
- Hellwig, Z. (1981). Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych [Multivariate comparative analysis and its application in the study of multivariate economic objects]. In: *Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną [Economic and mathematical methods and models in improving the management of the socialist economy]*. W., Welfe (ed.). Warszawa, PWE.
- Ilnicki, D. (2002). Próba określenia zmienności czasowej zjawisk przestrzennych metodami wzorcowymi – przykład metody Hellwiga [An attempt to determine the time variability of spatial phenomena by model methods – an example of Hellwig’s method]. In: *Możliwości i ograniczenia zastosowań metod badawczych w geografii społeczno-ekonomicznej i gospodarce przestrzennej [Possibilities and limitations of research methods application in social and economic geography and spatial economy]*. H., Rogacki (ed.). Poznań, Bogucki Wydawnictwo Naukowe.
- Jaworska, M., Luty, L. (2009). Ocena rozwoju społeczno-gospodarczego powiatów województwa małopolskiego [Evaluation of socio-economic development of poviats of Lesser Poland Voivodeship]. *Acta Scientiarum Polonorum Oeconomia*, 8(3), pp. 37–44.
- Jenks, G.F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, 7, pp. 186–190.
- Khamis, N., Sin, T.C., Hock, G.C. (2018). Segmentation of Residential Customer Load Profile in Peninsular Malaysia using Jenks Natural Breaks. *2018 IEEE 7th International Conference on Power and Energy, PECON 2018*, pp. 128–131. doi:10.1109/PECON.2018.8684113.
- Klopp, J.M., Petretta, D.L. (2017). The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities*, 63, pp. 92–97. doi:10.1016/j.cities.2016.12.019.
- Kłóska, R. (2017). Rozwój zrównoważony regionów w Polsce w ujęciu statystycznym [Sustainable development of regions in Poland in statistical perspective]. *Progress in Economic Sciences*, 4, pp. 159–176. doi:10.14595/PES/04/011.
- Koszel, M., Bartkowiak, P. (2018). Taksonomiczna miara zrównoważonego rozwoju obszarów metropolitalnych w Polsce [A taxonomic measure of sustainable development of metropolitan areas in Poland]. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie [Scientific Journals of Cracow University of Economics]*, 3(3(975)), pp. 83–100. doi:10.15678/znupek.2018.0975.0306.
- Kuciński, K. (2015). *Geografia ekonomiczna*. Warszawa, Oficyna Wolters Kluwer Business.
- Laskowska, I., Dańska-Borsiak, B. (2018). Analiza przestrzennego zróżnicowania rozwoju społecznego w Polsce na poziomie NUTS 3 z wykorzystaniem lokalnego indeksu HDI [Analysis of spatial diversity of social development in Poland at NUTS 3 level with the use of local HDI index]. *Acta Universitatis Lodziensis. Folia Oeconomica*, 1(333), pp. 111–131. doi:10.18778/0208-6018.333.08.
- Łogwiniuk, K. (2011). The use of taxonomic methods in the comparative analysis of the access to the ICT

- infrastructure by schoolchildren in Poland. *Economy and Management*, 1, pp. 7–23.
- Lu, D., Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5). doi:10.1080/01431160600746456.
- Łuczak, A., Kurzawa, I. (2017). Assessment of the Level of Sustainable Development in Districts in Poland Using Taxonomic Methods. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 469, pp. 109–118. doi:10.15611/pn.2017.469.11.
- Malina, A. (2020). Analiza przestrzennego zróżnicowania poziomu rozwoju społeczno-gospodarczego województw Polski w latach 2005–2017 [Analysis of spatial differentiation of the level of socio-economic development of Polish voivodeships in the years 2005–2017]. *Social Inequalities and Economic Growth*, 61(1), pp. 138–155. doi:10.15584/nsawg.2020.1.
- Matras, A. (2017). Ranking społeczno-gospodarczy jednostek samorządu terytorialnego z uwzględnieniem poziomu zadłużenia oprocentowanego per capita [Socio-economic ranking of local government units taking into account the level of interest-bearing debt per capita]. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach [Economic Studies. Scientific Journals of the University of Economics in Katowice]*, 341, pp. 230–239.
- Michoń, D. (2017). Zróżnicowanie rozwoju społeczno-gospodarczego województw ze względu na realizację celów polityki spójności [Differentiation of socio-economic development of voivodeships in relation to the implementation of cohesion policy objectives]. *Wiadomości Statystyczne [Statistical News]*, 12(679), pp. 80–94.
- Mikuła, A. (2020). Level of Sustainable Development of Counties in Poland. *Annals of the Polish Association of Agricultural and Agribusiness Economists*, XXII(1), pp. 241–249. doi:10.5604/01.3001.0013.7792.
- Misiewicz, E.I., Roszkowska, E., Rogowski, J. (2019). Ocena poziomu ładu gospodarczego województw Polski w 2015 r. [Assessment of the level of economic governance of Poland's provinces in 2015]. *Optimum. Economic Studies*, 3(3(97)), pp. 3–18. doi:10.15290/oes.2019.03.97.01.
- Moran, D.D., Wackernagel, M., Kitzes, J.A., Goldfinger, S.H., Boutaud, A. (2008). Measuring sustainable development – Nation by nation. *Ecological Economics*, 64(3). doi:10.1016/j.ecolecon.2007.08.017.
- Moreno Pires, S., Fidéls, T., Ramos, T.B. (2014). Measuring and comparing local sustainable development through common indicators: Constraints and achievements in practice. *Cities*, 39. doi:10.1016/j.cities.2014.02.003.
- Mori, K., Christodoulou, A. (2012). Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review*, 32(1), pp. 94–106. doi:10.1016/j.eiar.2011.06.001.
- Mortensen, L.F. (2013). Measuring sustainable development. In: *Environmental Management in Practice: Instruments for Environmental Management*. P., Compton, D., Devuyt, L., Hens, B., Nath (eds.). Londyn, Routledge.
- Morton, S., Pencheon, D., Squires, N. (2017). Sustainable Development Goals (SDGs), and their implementation. *British Medical Bulletin*, 124. doi:10.1093/bmb/ldx031.
- Niemczyk, A. (2001). Poziom rozwoju społecznego w nowym układzie administracyjnym Polski [The level of social development in the new administrative arrangement of Poland]. *Przegląd Statystyczny [Statistical Review]*, 48(3–4), pp. 289–300.
- Nourry, M. (2008). Measuring sustainable development: Some empirical evidence for France from eight alternative indicators. *Ecological Economics*, 67(3). doi:10.1016/j.ecolecon.2007.12.019.
- Nowak, E. (1990). *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych [Taxonomic methods in the classification of socio-economic objects]*. Warszawa, PWE.
- Pearce, D., Hamilton, K., Atkinson, G. (1996). Measuring sustainable development: progress on indicators. *Environment and Development Economics*, 1(1). doi:10.1017/s1355770x00000395.
- Podogrodzka, M. (2011). Analiza zjawisk społeczno-ekonomicznych z zastosowaniem metod taksonomicznych [Analysis of socio-economic phenomena using taxonomic methods]. *Wiadomości Statystyczne [Statistical News]*, 11, pp. 26–41.
- Podstawka, M., Suchodolski, B. (2018). Assessment of the level of economic and social development of regions using the Hellwig taxonomic development measure. *Proceedings of the 2018 VII International Scientific Conference Determinants of Regional Development*, 1, pp. 187–201. doi:10.14595/CP/01/013.
- Polish National Register of Borders. (2020). <http://www.gugik.gov.pl/pzgik/dane-bez-oplat/dane-z-panstwo->

- wego-rejestru-granic-i-powierzchni-jednostek-podzialow-terytorialnych-kraju-prg, date: 16.07.2020.
- Pomianek, I. (2010). Poziom rozwoju społeczno-gospodarczego obszarów wiejskich województwa warmińsko-mazurskiego [Level of Socio-Economic Development of Rural Areas in Warmińsko-Mazurskie Voivodeship]. *Acta Scientiarum Polonorum Oeconomia*, 9, pp. 227–239.
- Prus, B., Król, K. (2017). Ocena zastosowania wybranych metod taksonomicznych do klasyfikacji zjawisk społeczno-gospodarczych [Evaluation of the application of selected taxonomic methods for classifying socio-economic phenomena]. *Acta Scientiarum Polonorum Formatio Circumiectus*, 16(2), pp. 179–197. doi:10.15576/ASP.FC/2017.16.2.179.
- Rzasa, K., Ogryzek, M., Żróbek, R. (2019). The land transfer from the state treasury to local government units as a factor of social development of rural areas in Poland. *Land*, 8(11). doi:10.3390/land8110170.
- Salamon, J. (2005). Badania wielofunkcyjnego rozwoju obszarów wiejskich województwa świętokrzyskiego [Studies on multifunctional development of rural areas in Świętokrzyskie voivodeship]. *Infrastruktura i Ekologia Terenów Wiejskich [Infrastructure and Ecology of Rural Areas]*, 4, pp. 145–155.
- Schleicher-Tappeser, R. (2018). Sustainable development in the European Union – Monitoring Report on Progress Towards the SDGS in an EU Context. *Greener Management International*, 2001(36), pp. 50–66. doi:10.9774/gleaf.3062.2001.wi.00007.
- Schober, P., Boer, C., Schwarte, L.A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, 126(5), pp. 1763–1768. doi:10.1213/ANE.0000000000002864.
- Senetra, A., Szarek-Iwaniuk, P. (2019). The implementation of the Wrocław taxonomic method for the identification and evaluation of problem areas in the Warmia and Mazury Region in Poland – A case study. *Socio-Economic Planning Sciences*, 67, pp. 43–57. doi:10.1016/j.seps.2018.09.003.
- Senetra, A., Szarek-Iwaniuk, P. (2020). Socio-economic development of small towns in the Polish Citaslow Network – A case study. *Cities*, 103, p. 102758. doi:10.1016/j.cities.2020.102758.
- Sharma, A., Tewari, R., Rana, S.S., Soni, R., Soni, S.K. (2016). Cellulases: Classification, Methods of Determination and Industrial Applications. *Applied Biochemistry and Biotechnology*, 179. doi:10.1007/s12010-016-2070-3.
- Shen, L.Y., Jorge Ochoa, J., Shah, M.N., Zhang, X. (2011). The application of urban sustainability indicators – A comparison between various practices. *Habitat International*, 35(1), pp. 17–29. doi:10.1016/j.habitatint.2010.03.006.
- Sołek, K., Sowa, B. (2019). Diversification of the Social Development of Podkarpackie Province Communes. *Economic and Regional Studies / Studia Ekonomiczne i Regionalne*, 12(1), pp. 45–55. doi:10.2478/ers-2019-0005.
- Spangenberg, J.H. (2015). Indicators for sustainable development. In: *Routledge International Handbook of Sustainable Development*. M., Redclift, D., Springett (eds.). Londyn, Routledge.
- Stafford-Smith, M., Griggs, D., Gaffney, O., Ullah, F., Reyers, B., Kanie, N., Stigson, B., Shrivastava, P., Leach, M., O'Connell, D. (2017). Integration: the key to implementing the Sustainable Development Goals. *Sustainability Science*, 12. doi:10.1007/s11625-016-0383-3.
- Stec, M. (2012). Analiza porównawcza rozwoju społeczno-gospodarczego powiatów województwa podkarpackiego [A comparative analysis of socio-economic development of powiats in Podkarpackie voivodeship]. *Nierówności Społeczne a Wzrost Gospodarczy [Social Inequality and Economic Growth]*, 25, pp. 180–190.
- Sustainable Development Indicators. (2017). wskaźniki-zrp.stat.gov.pl, date 15.07.2020.
- Tanguay, G.A., Rajaonson, J., Lefebvre, J.F., Lanoie, P. (2010). Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecological Indicators*, 10(2). doi:10.1016/j.ecolind.2009.07.013.
- Tarka, D. (2010). Własności cech diagnostycznych w badaniach typu taksonomicznego [Properties of diagnostic features in taxonomic type studies]. *Ekonomia i Zarządzanie [Economics and Management]*, 2(4), pp. 194–203.
- Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*. doi:10.1016/j.aci.2018.08.003.
- Toloie-Eshlaghy, A., Chitsaz, S., Karimian, L., Clarkchi, R. (2011). A classification of qualitative research methods. *Research Journal of International Studies*, 20, pp. 106–123.
- United Nations. (2007). *Indicators of Sustainable Development: Guidelines and Methodologies*. New York, United Nations.
- United Nations. (2015). *Sustainable Development Knowledge Platform*. <https://sustainabledevelopment.un.org/index.html>, date: 16.07.2020.



Urząd Statystyczny w Katowicach. (2018). *Produkt krajowy brutto i wartość dodana brutto w przekroju regionów w 2016 r.* [Gross domestic product and gross value added by region in 2016]. [https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5482/7/1/1/pkb\\_i\\_wartosc\\_dodana\\_brutto\\_w\\_przekroju\\_regionow\\_w\\_2016.pdf](https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5482/7/1/1/pkb_i_wartosc_dodana_brutto_w_przekroju_regionow_w_2016.pdf), date: 16.07.2020.

World Commission on Environment and Development. (1987). *Our Common Future*. New York, Oxford University Press.

Ziemiańczyk, U. (2010). Ocena poziomu rozwoju społeczno-gospodarczego gmin wiejskich i miejsko-wiejskich w województwie małopolskim [Evaluation of the level of socio-economic development of rural and urban-rural communes in the Małopolskie Voivodeship]. *Infrastruktura i Ekologia Terenów Wiejskich* [Infrastructure and Ecology of Rural Areas], 14, pp. 31–40.

