

# PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

# RESEARCH PAPERS

of Wrocław University of Economics

Nr 331

## Problemy rozwoju regionalnego i lokalnego

Redaktorzy naukowci

Elżbieta Sobczak, Beata Bał-Domańska,  
Marek Obrębalski



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu  
Wrocław 2014

Redaktor Wydawnictwa: Aleksandra Śliwka  
Redaktor techniczny: Barbara Łopusiewicz  
Korektor: Barbara Cibis  
Łamanie: Małgorzata Czupryńska  
Projekt okładki: Beata Dębska

Projekt współfinansowany z budżetu województwa dolnośląskiego



**DOLNY  
ŚLĄSK**

Publikacja jest dostępna w Internecie na stronach:  
[www.ibuk.pl](http://www.ibuk.pl), [www.ebscohost.com](http://www.ebscohost.com),  
w Dolnośląskiej Bibliotece Cyfrowej [www.dbc.wroc.pl](http://www.dbc.wroc.pl),  
The Central and Eastern European Online Library [www.ceeol.com](http://www.ceeol.com),  
a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon  
[http://kangur.uek.krakow.pl/bazy\\_ae/bazekon/nowy/index.php](http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php)

Informacje o naborze artykułów i zasadach recenzowania znajdują się  
na stronie internetowej Wydawnictwa  
[www.wydawnictwo.ue.wroc.pl](http://www.wydawnictwo.ue.wroc.pl)

Kopiowanie i powielanie w jakiegokolwiek formie  
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu  
Wrocław 2014

**ISSN 1899-3192**  
**ISBN 978-83-7695-456-1**

Wersja pierwotna: publikacja drukowana

Druk i oprawa:  
EXPOL, P. Rybiński, J. Dąbek, sp.j.  
ul. Brzeska 4, 87-800 Włocławek

## Spis treści

Wstęp.....	9
<b>Beata Bal-Domańska, Michał Bernard Pietrzak:</b> Modelowanie wzrostu gospodarczego na podstawie rozszerzonego modelu Solowa-Swana z uwzględnieniem aspektu przestrzennego.....	11
<b>Grażyna Bojęć:</b> Nowy wskaźnik zadłużenia a koszty obsługi długu w jednostkach samorządu terytorialnego na przykładzie powiatu jeleniogórskiego.....	19
<b>Dariusz Głuszczyk:</b> Kredyty bankowe jako źródło finansowania działalności innowacyjnej przedsiębiorstw – analiza w przekroju regionów Polski.....	30
<b>Dariusz Głuszczyk:</b> Kredyt technologiczny jako instrument wsparcia innowacji małych i średnich przedsiębiorstw – analiza w przekroju regionów Polski.....	41
<b>Małgorzata Januszewska, Elżbieta Nawrocka:</b> Zmiany czynników lokalizacji podmiotów turystycznych .....	53
<b>Marek Kiczek:</b> Zmiany udziału dochodów własnych w dochodach ogółem gmin województwa podkarpackiego w latach 2006, 2012.....	64
<b>Renata Lisowska:</b> Wsparcie rozwoju małych i średnich przedsiębiorstw przez samorząd terytorialny w obszarach zmarginalizowanych.....	75
<b>Olga Ławińska:</b> Ocena efektywności inwestycji współfinansowanych funduszami Unii Europejskiej na przykładzie budowy oczyszczalni ścieków i kanalizacji sanitarnej w gminie Kłomnice w latach 2009-2012.....	85
<b>Marek Obrębalski, Marek Walesiak:</b> Terytorialny wymiar polityki rozwoju regionalnego województwa dolnośląskiego w latach 2014-2020 .....	96
<b>Katarzyna Przybyła:</b> Poziom rozwoju infrastruktury technicznej w miastach wojewódzkich Polski.....	106
<b>Adam Przybyłowski:</b> Gospodarka regionalna w aspekcie pomiaru zrównoważonego transportu.....	116
<b>Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska:</b> Wykorzystanie analizy wielogrupowej do porównania rynku pracy w regionach.....	125
<b>Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska:</b> Sposoby wyznaczania środków regionów na potrzeby analiz przestrzennych.....	134
<b>Alicja Sekuła, Beata A. Basińska:</b> Dlaczego subwencje nie są rozwojowe? Próba identyfikacji przyczyn braku wpływu subwencji na wydatki inwestycyjne	146
<b>Elżbieta Sobczak:</b> Harmonijność inteligentnego rozwoju województw Polski .....	158
<b>Roman Sobczak:</b> Zróżnicowanie zasobów ludzkich w nauce i technice w krajach Unii Europejskiej.....	169

<b>Wioleta Sobczak, Lilianna Jabłońska, Lidia Gunerka:</b> Zmiany strukturalne w powierzchni gruntów użytkowanych ogrodniczo w województwie mazowieckim w świetle spisów rolnych.....	180
<b>Danuta Strahl, Andrzej Sokółowski:</b> Propozycja podejścia metodologicznego do oceny zależności między inteligentnym rozwojem a wrażliwością na kryzys ekonomiczny w wymiarze regionalnym .....	190
<b>Agnieszka Stacherzak, Maria Heldak, Jan Kazak:</b> Obciążenia finansowe gmin kosztami realizacji dróg .....	201
<b>Artur Stec:</b> Związek między funkcją turystyczną a wydatkami na turystykę w miastach na prawach powiatu w województwie podkarpackim w latach 2008-2012.....	213
<b>Aldona Standar:</b> Rozwój infrastruktury wodno-kanalizacyjnej na obszarach wiejskich województwa wielkopolskiego po wstąpieniu Polski do Unii Europejskiej.....	224
<b>Justyna Weltrowska, Wojciech Kisiała:</b> Obszary koncentracji ubóstwa w strukturze przestrzennej miasta (na przykładzie Poznania).....	235
<b>Wioletta Wierzbicka:</b> Potencjał innowacyjny polskich regionów – analiza taksonomiczna.....	246
<b>Justyna Wilk:</b> Dane symboliczne w analizie regionalnego zróżnicowania sytuacji gospodarczej .....	257
<b>Dariusz Zawada:</b> Identyfikacja i ocena walorów użytkowych miast – studium przypadku dla Jeleniej Góry i Legnicy.....	270
<b>Marcelina Zapotoczna, Joanna Cymerman:</b> Zastosowanie analizy wielowymiarowej do oceny rozwoju lokalnych rynków nieruchomości mieszkaniowych na przykładzie miast wojewódzkich.....	282

## Summaries

<b>Beata Bal-Domańska, Michał Bernard Pietrzak:</b> Economic growth modelling based on the augmented Solow-Swan model considering the special aspect ..	18
<b>Grażyna Bojęć:</b> New debt indicator vs. debt servicing costs in self-government units: Jelenia Góra county example.....	29
<b>Dariusz Głuszczyk:</b> Bank credits as a source of financing innovative activities of enterprises – an analysis by regions of Poland.....	40
<b>Dariusz Głuszczyk:</b> Technology credit as an instrument of support to small and medium-sized enterprises – an analysis by regions of Poland.....	52
<b>Małgorzata Januszewska, Elżbieta Nawrocka:</b> Changes in factors of tourism entities location .....	63
<b>Marek Kiczek:</b> Changes of the participation level of own communes income in the total income of Podkarpackie Voivodeship communes in 2006, 2012.....	74
<b>Renata Lisowska:</b> Support for the development of small and medium-sized enterprises in marginalised areas provided by local government .....	84

<b>Olga Ławińska:</b> Effectiveness evaluation of co-financed European Union funds investment on the example of sewage treatment plant and sewage system in Kłomnice community in the years 2009-2012 .....	95
<b>Marek Obrębalski, Marek Walesiak:</b> Territorial dimension of regional development policy in Lower Silesia region in 2014-2020 .....	105
<b>Katarzyna Przybyła:</b> The level of technical infrastructure in Voivodeship cities in Poland .....	115
<b>Adam Przybyłowski:</b> Regional economy in the context of sustainable transport measurement .....	124
<b>Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska:</b> The application of multiple group analysis in labour market analysis of regions .....	133
<b>Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska:</b> The ways of outlining the centers of regions for the purposes of spatial analyses .....	145
<b>Alicja Sekuła, Beata A. Basińska:</b> Why are not subsidies developmental? An attempt to identify the reasons of the lack of influence on investment expenditures .....	157
<b>Elżbieta Sobczak:</b> Harmonious smart growth of voivodeships in Poland .....	168
<b>Roman Sobczak:</b> Diversity of human resources in science and technology in the European Union countries .....	179
<b>Wioleta Sobczak, Lilianna Jabłońska, Lidia Gunerka:</b> Structural changes in horticultural production in the Mazovian Voivodeship in the light of the national agricultural census .....	189
<b>Danuta Strahl, Andrzej Sokółowski:</b> The proposal of methodological approach to the assessment of relations between smart growth and vulnerability to economic crisis at the regional level .....	200
<b>Agnieszka Stacherzak, Maria Heldak, Jan Kazak:</b> Financial burden of municipalities with the costs of roads development .....	212
<b>Artur Stec:</b> The relationship between tourist function and expenditure on tourism in cities with county rights in the Podkarpackie Voivodeship in 2008-2012 .....	222
<b>Aldona Standar:</b> The development of water supply and sewerage system in rural areas of the Great Poland Voivodeship after Polish accession to the European Union .....	234
<b>Justyna Weltrowska, Wojciech Kisiała:</b> Areas of concentration of poverty in the city's spatial structure (the case study of Poznań) .....	245
<b>Wioletta Wierzbicka:</b> Innovative potential of Polish regions – taxonomic analysis .....	256
<b>Justyna Wilk:</b> Symbolic data in the analysis of regional diversification of economic situation .....	269
<b>Dariusz Zawada:</b> Identification and assessment of utility values of the cities – case study of Jelenia Góra and Legnica .....	281
<b>Marcelina Zapotoczna, Joanna Cymerman:</b> Applying multidimensional analysis to assess the development of local housing property markets on the basis of voivodeship cities .....	293

**Justyna Wilk**

Uniwersytet Ekonomiczny we Wrocławiu

---

## DANE SYMBOLICZNE W ANALIZIE REGIONALNEGO ZRÓŻNICOWANIA SYTUACJI GOSPODARCZEJ

---

**Streszczenie:** Dane symboliczne znajdują zastosowanie w przestrzennej analizie sytuacji gospodarczej. Umożliwiają uwzględnienie struktury zjawiska w obrębie porównywanych regionów, szczególnie w przypadku znacznych dysproporcji rozwojowych. Celem artykułu jest przedstawienie koncepcji danych symbolicznych i możliwości ich wykorzystania w opisie sytuacji gospodarczej jednostek terytorialnych. W artykule wskazano problemy pomiaru sytuacji gospodarczej w badaniach porównawczych jednostek terytorialnych. Zaprezentowano koncepcję i konstrukcję danych symbolicznych w opisie sytuacji jednostek terytorialnych, a także metody ich analizy. Przedstawiono również przykład empiryczny dotyczący wykorzystania danych symbolicznych w analizie poziomu zróżnicowania sytuacji gospodarczej w Polsce.

**Słowa kluczowe:** dane symboliczne, odległość ekonomiczna, zróżnicowanie regionalne.

DOI: 10.15611/pn.2014.331.24

### 1. Wstęp

Analiza sytuacji gospodarczej jednostki terytorialnej (np. powiatu, regionu, kraju, Unii Europejskiej itd.) jest zagadnieniem złożonym i wielowymiarowym. Wymaga uwzględnienia wielu powiązanych i wzajemnie warunkujących się czynników. O sytuacji gospodarczej stanowi profil gospodarczy, poziom innowacyjności, poziom inwestycji, sytuacja rynku pracy, dochodowość przemysłu, kondycja finansowa przedsiębiorstw, napływ kapitału zagranicznego itd. (zob. np. [Strahl (red.), 2006; Bal-Domańska, Wilk 2011]).

Ocena sytuacji gospodarczej jednostki terytorialnej jest wypadkową sytuacji jej podregionów. Nieuwzględnienie terytorialnej struktury zjawiska może prowadzić do błędnej interpretacji wyników badania, szczególnie w przypadku analiz porównawczych. Jednym z rozwiązań tego problemu jest zastosowanie w pomiarze podobieństwa regionów miar odległości, które w swojej konstrukcji uwzględniają strukturę zjawisk (zob. np. [Strahl 2006; 2007]). Alternatywną propozycją, szczególnie w sytuacji, gdy analiza porównawcza regionów nie bazuje na pomiarze odległości, jest zastosowanie metodologii analizy danych symbolicznych (zob. [Bock, Diday (red.) 2000; Billard, Diday 2006; Diday, Noirhomme-Fraiture (red.) 2008]).

Celem artykułu jest propozycja uwzględnienia danych symbolicznych w reprezentacji sytuacji gospodarczej jednostek terytorialnych oraz zastosowanie tego podejścia w analizie regionalnego zróżnicowania sytuacji gospodarczej w Polsce w 2010 r. W pierwszej części artykułu rozważone zostaną problemy związane z pomiarem sytuacji gospodarczej w analizach porównawczych jednostek terytorialnych. W kolejnych częściach pracy zaprezentowana zostanie koncepcja i ich konstrukcja danych symbolicznych w opisie sytuacji jednostek terytorialnych, a także metody ich analizy. W ostatniej części przedstawiony zostanie przykład empiryczny dotyczący wykorzystania danych symbolicznych w analizie poziomu zróżnicowania sytuacji gospodarczej w Polsce.

## **2. Problemy pomiaru w analizie regionalnego zróżnicowania sytuacji gospodarczej**

Pomiar sytuacji gospodarczej wymaga rozważenia szeregu problemów natury metodologicznej. Jedną z trudności w analizach regionalnych jest liczny zbiór danych. Badaniu podlega zazwyczaj wiele obiektów (jednostek terytorialnych), które są opisane za pomocą wielu zmiennych (charakterystyk). Taka sytuacja ma miejsce, gdy rozpatrujemy np. czynniki warunkujące społeczno-gospodarczy rozwój regionów klasy NTS-2 w UE.

Ponadto uwzględnienie wyższego poziomu podziału terytorialnego prowadzi najczęściej to zmniejszenia precyzji opisu zjawiska. Jest to spowodowane stosowaniem w opisie jednostek terytorialnych wyższego szczebla wartości agregatowych lub statystyk opisowych, np. średniej arytmetycznej, mediany. Takie podejście może skutkować błędną interpretacją wyników badania, związaną z rozpatrywaniem zjawiska bez uwzględnienia jego terytorialnej struktury. Na przykład stopa bezrobocia rejestrowanego w województwie dolnośląskim w 2012 r. wynosiła 13,5%. Natomiast w powiatach województwa jej wartość mieściła się w przedziale od 5,6% do 27,5%, przy czym  $\frac{3}{4}$  powiatów województwa uzyskało znacznie wyższe udziały procentowe od średniej wojewódzkiej. Rozwiązaniem w tej sytuacji jest konstrukcja danych symbolicznych (zob. [Wilk 2011; 2012b]).

## **3. Konstrukcja danych symbolicznych w opisie sytuacji jednostek terytorialnych**

Zastosowanie danych symbolicznych umożliwia opisanie obiektów (jednostek terytorialnych) za pomocą zmiennych o realizacjach w postaci przedziałów wartości, zbiorów kategorii, struktur udziałowych. Daje również możliwość uwzględnienia zależności zmiennych powiązanych relacjami. Rodzaje zmiennych symbolicznych zaprezentowano w tab. 1.

Pozwala także scharakteryzować jednostki terytorialne szczebla wyższego (np. województwa NTS-2) na podstawie sytuacji znajdujących się w ich obrębie jedno-

**Tabela 1.** Rodzaje zmiennych symbolicznych

Rodzaj zmiennej symbolicznej	Realizacje zmiennej	Przykłady
Przedziałowa ( <i>interval-valued variable</i> )	przedziały wartości	przedziały wieku lub dochodu respondentów; przedziały cenowe produktów
Wielowariantowa ( <i>multi-valued variable</i> )	zbiory kategorii	posiadane kategorie prawa jazdy; składniki produktu; znajomość języków obcych
Wielowariantowa z wagami ( <i>modal variable</i> )	struktury udziałowe	struktura wydatków (na wyżywienie, odzież, usługi itd.); udziały procentowe ludności według ekonomicznych grup wieku
Strukturalna ( <i>dependent variable</i> )	struktura hierarchiczna, logiczna, taksonomiczna	modele i marki samochodów; wzrost i waga dzieci; taksonomia regionów geograficznych

Źródło: opracowanie własne z wykorzystaniem [Bock, Diday (red.) 2000].

stek szczebla niższego (np. powiatów NTS-4). Obiekty rozumiane jako elementarna jednostka badania, opisane zmiennymi symbolicznymi, nazywane są obiektami symbolicznymi I rzędu. Natomiast obiekty nadrzędne powstałe w wyniku agregacji danych w ujęciu klasycznym (pojedyncza kategoria lub liczba rzeczywista) określane są jako obiekty symboliczne II rzędu.

**Tabela 2.** Rodzaje obiektów symbolicznych

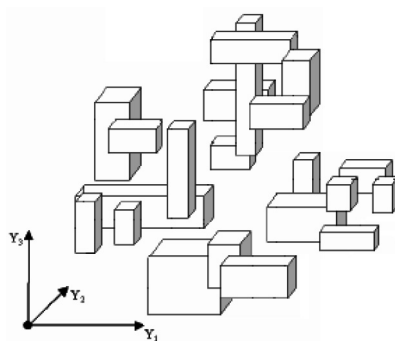
Rodzaj obiektu	Specyfika	Przykłady
Obiekt symboliczny I rzędu ( <i>first order symbolic object</i> )	Obiekt rozumiany w sensie klasycznym (elementarna jednostka badania)	respondent; kraj; produkt
Obiekt symboliczny II rzędu ( <i>second order symbolic object</i> )	Obiekty będące wynikiem agregacji obiektów symbolicznych I rzędu	region złożony z podregionów; model samochodu (np. Volkswagen Golf)

Źródło: opracowanie własne z wykorzystaniem [Bock, Diday (red.) 2000].

Zbiór obiektów symbolicznych, opisanych za pomocą trzech zmiennych symbolicznych, o realizacjach w postaci przedziałów liczbowych zaprezentowano na rys. 1.

Dane symboliczne otrzymuje się w wyniku agregacji obiektów lub zmiennych (zob. [Bock, Diday (red.) 2000; Diday, Noirhomme-Fraiture (red.) 2008; Billard, Diday 2006; Wilk 2012]). Zmienną symboliczną o realizacjach w postaci przedziałów liczbowych otrzymuje się poprzez wyznaczenie minimalnej i maksymalnej wartości zmiennej obserwowanej w jednostkach szczebla niższego bądź z wykorzystaniem kwartyli (zob. przykład w tab. 3).





**Rys. 1.** Zbiór obiektów symbolicznych w przestrzeni trójwymiarowej

Źródło: opracowanie własne.

**Tabela 3.** Przeciętne miesięczne wynagrodzenie w Polsce w 2010 r. (zł)

NTS-1 (region)		NTS-3 (subregion)		Statystyki opisowe dla NTS-3		Zmienna przedziałowa dla NTS-1
Nazwa	wartość	nazwa	wartość	minimum	maksimum	
Centralny	4025,26	łódzki	2705,93	2666,84	4694,47	[2666,84; 4694,47]
		city of Łódź	3243,15			
		piotrkowski	3297,76			
		sieradzki	2666,84			
		skierniewicki	2810,76			
		ciechanowsko-płocki	3366,54			
		ostrołęcko-siedlecki	2950,04			
		radomski	3054,40			
		Warszawa	4694,47			
		warszawski wschodni	3209,11			
		warszawski zachodni	3657,19			
Wschodni	2987,67	białski	2841,15	2650,41	3410,28	[2650,41; 3410,28]
		...	.			
		sandomiersko-jędrzejowski	2982,23			
...	...	...	...	...	...	...

Źródło: opracowanie własne na podstawie danych BDL GUS.

Zmienną wielowariantową z wagami otrzymuje się na podstawie określenia kategorii zmiennej (np. niski, średni i wysoki poziom bezrobocia) i wyznaczenia odsetka jednostek (np. LAU 1) w każdej kategorii, np. NTS-2 (zob. przykład w tab. 4).

**Tabela 4.** Stopa bezrobocia rejestrowanego w Polsce w 2010 r. (%)

NTS-2 (województwo)		NTS-4 (powiat)		Odsetek powiatów (NTS-4) spełniający kategorię*			Zmienna wielowariantowa z wagami dla NTS-2
Nazwa	wartość	nazwa	wartość	niskie	średnie	wysokie	
Mazowieckie	9,7	m. Warszawa	3,5	0,2	0,5	0,3	{niskie (0,2), średnie (0,5), wysokie (0,3)}
		warszawski zachodni	5,9				
		.	.				
		.	.				
		radomski	30,8				
		sztywnowiecki	36,0				
Kujawsko-pomorskie	17,0	Bydgoszcz	8,0	0,1	0,3	0,6	{niskie (0,1), średnie (0,3), wysokie (0,6)}
		lipnowski	28,9				
...	...	...	...	...	...	...	...

\* Niskie bezrobocie (stopa bezrobocia poniżej 10%), średnie (10-20%), wysokie (powyżej 20%).

Źródło: opracowanie własne na podstawie danych BDL GUS.

**Tabela 5.** Ludność według ekonomicznych grup wieku w Polsce w 2010 r. (%)

NTS-5 (miasta)	Odsetek ludności według ekonomicznych grup wieku*			Zmienna wielowariantowa z wagami (NTS-5)
	przedprodukcyjny	produkcyjny	poprodukcyjny	
Łódź	13,9	64,3	21,8	{przedprodukcyjny (0,139), produkcyjny (0,643), poprodukcyjny (0,218)}
Warszawa	15,0	63,9	21,1	{przedprodukcyjny (0,150), produkcyjny (0,639), poprodukcyjny (0,211)}
Kraków	15,5	65,2	19,3	{przedprodukcyjny (0,155), produkcyjny (0,652), poprodukcyjny (0,193)}
Wrocław	14,7	65,7	19,6	{przedprodukcyjny (0,147), produkcyjny (0,657), poprodukcyjny (0,196)}
...	...	...	...	...

\* Wiek przedprodukcyjny (17 lat i mniej), produkcyjny (mężczyźni 18-64, kobiety 18-59), poprodukcyjny (mężczyźni 65 i więcej, kobiety 60 i więcej)

Źródło: opracowanie własne na podstawie danych BDL GUS.

Dane symboliczne mogą być także rezultatem łączenia kategorii (np. rodzaju obiektów sportowych, ekonomicznych grup wieku) opisujących jednostki teryto-

rialne. W tabeli 5 zaprezentowano sposób otrzymywania zmiennej o realizacjach w postaci struktury udziałowej. W pierwszej kolejności zdefiniowano kategorie zmiennej, a następnie wyznaczono odsetki jednostek terytorialnych spełniających poszczególne kategorie.

**Tabela 6.** Szkoły wyższe według rodzaju w 2010 r. (szt.)

NTS-2	Liczba szkół wyższych (uniwersytetów, politechnik, akademii) według rodzaju					Lista kategorii dla NTS-2
	techniczne	rolnicze	pedagogiczne	medyczne	sztuk pięknych	
Kujawsko- pomorskie	1	1	0	0	1	{techniczne, rolnicze, sztuk pięknych}
Warmińsko- mazurskie	0	0	2	0	0	{pedagogiczne}
Śląskie	4	0	2	1	2	{techniczne, pedagogiczne, medyczne, sztuk pięknych}
Dolnośląskie	1	1	1	1	2	{techniczne, rolnicze, pedagogiczne, medyczne, sztuk pięknych}
Podkarpackie	1	0	0	0	0	{techniczne}
...	...	...	...	...	...	...

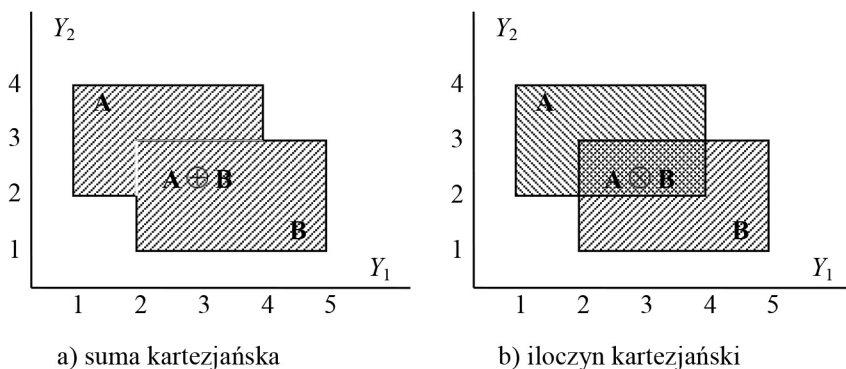
Źródło: opracowanie własne na podstawie danych BDL GUS.

Z kolei w tabeli 6 zaprezentowano sposób definiowania zmiennej wielowariantowej. Zmienną symboliczną tego rodzaju tworzy się, określając, które kategorie realizowane są przez jednostkę terytorialną.

#### **4. Metody analizy danych symbolicznych w ocenie poziomu zróżnicowania regionalnego**

Podobieństwo jednostek terytorialnych opisanych za pomocą danych symbolicznych można określić z wykorzystaniem miar odległości wypracowanych na gruncie analizy danych symbolicznych (zob. np. [Bock, Diday (red.) 2000, s. 153-185]). W swojej konstrukcji wykorzystują one pojęcie sumy i iloczynu kartezjańskiego, których ilustrację przedstawiono na rys. 2.

Miary odległości obiektów opisanych zmiennymi o realizacjach w postaci przedziałów liczbowych, zbiorów kategorii i zmiennych zależnych zaproponowane zostały przez Ichino i Yaguchiego, de Carvalho, Gowdę i Didaya (zob. [Malerba i in. 2001; Wilk 2006a]). Miary odległości obiektów opisanych zmiennymi w postaci



**Rys. 2.** Ilustracja sumy i iloczynu kartezjańskiego

Źródło: opracowanie własne.

struktur udziałowych są w większości adaptacją miar sformułowanych dla rozkładów prawdopodobieństwa, klasycznych miar odległości obiektów, a także miar wykorzystywanych w segmentacji obrazów [Malerba, Esposito, Monopoli 2002, s. 33-35; Wilk 2006b].

Miara odległości obiektów względem  $k$ -tej zmiennej zaproponowana przez Ichino i Yaguchiego jest zapisywana wzorem [Ichino, Yaguchi 1994]:

$$d_{ijk} = \mu(v_{ik} \oplus v_{jk}) - \mu(v_{ik} \otimes v_{jk}) + \gamma \nu(v_{ik}, v_{jk}), \quad (1)$$

gdzie:  $i, j$  – numer obiektu symbolicznego,  $k$  – numer zmiennej symbolicznej,

$$\nu(v_{ik}, v_{jk}) \equiv 2\mu(v_{ik} \otimes v_{jk}) - \mu(v_{ik}) - \mu(v_{jk}),$$

$\gamma$  – parametr,  $\gamma \in [0; 0,5)$ ,

$v_{ik} \oplus v_{jk} \equiv \{\min\{\underline{v}_{ik}, \underline{v}_{jk}\}, \max\{\overline{v}_{ik}, \overline{v}_{jk}\}\}$  dla zmiennej w postaci przedziału liczbowego,

$\underline{v}_{ik}, \underline{v}_{jk}$  ( $\overline{v}_{ik}, \overline{v}_{jk}$ ) – dolne (górne) końce przedziału  $k$ -tej zmiennej, odpowiednio dla obiektów  $i$  oraz  $j$ ,

$$v_{ik} \otimes v_{jk} \equiv v_{ik} \cap v_{jk}.$$

Miara odległości obiektów względem zbioru zmiennych  $Y$  jest metryką Minkowskiego [Ichino, Yaguchi 1994]:

$$d_{ij} = \left[ \sum_{k=1}^p (\omega_k d_{ijk})^\lambda \right]^{1/\lambda}, \quad (2)$$

gdzie  $\omega_k$  – waga  $k$ -tej zmiennej ( $k = 1, p$ ),  $\lambda \geq 1$ .

**Tabela 7.** Metody wielowymiarowej analizy danych symbolicznych

Nazwa metod	Rodzaj	Przykładowe algorytmy
Drzewa decyzyjne	drzewa klasyfikacyjne	TREE, BDT, SDT
	drzewa regresyjne	–
Skalowanie wielowymiarowe	metody bazujące na macierzy odległości	bi-dimensional-mapping
	metody bazujące na tablicy danych symbolicznych	InterScal, SymScal, I-Scal
Analiza skupień	metody hierarchiczne	Warda, kompletnego połączenia, Brito, Chavent
	metody niehierarchiczne	<i>k</i> -medoidów, SCLUST, DCLUST

Źródło: opracowanie własne na podstawie prac [Diday, Noirhomme-Fraiture (red.) 2008; Gatnar, Walesiak (red.) 2011].

W analizie regionalnego zróżnicowania sytuacji gospodarczej zastosowanie mają przede wszystkim metody wielowymiarowej analizy statystycznej, takie jak np. analiza skupień czy skalowanie wielowymiarowe (zob. np. [Hair i in. 2006; Everitt, Dunn 2001]). W przypadku kiedy jednostki terytorialne opisane są danymi symbolicznymi, należy zastosować metody bazujące na macierzy odległości bądź tablicy danych symbolicznych (zob. np. [Diday, Noirhomme-Fraiture (red.) 2008; Wilk 2010; Gatnar, Walesiak (red.) 2011]). Przykładowe algorytmy zaprezentowano w tab. 7.

W analizie danych symbolicznych zastosowanie mają programy statystyczne, takie jak SODAS oraz R, w tym pakiety symbolicDA (Dudek, Pełka, Wilk), clusterSim (Walesiak, Dudek) i RSDA (Rodriguez). Przegląd i porównanie dostępnego oprogramowania zaprezentowano w pracy [Wilk 2012a].

## 5. Analiza poziomu zróżnicowania sytuacji gospodarczej na podstawie danych symbolicznych

Celem badania jest ocena poziomu regionalnego zróżnicowania sytuacji gospodarczej w Polsce. Analizą objęto sytuację gospodarczą 16 województw w 2010 r. Uwzględniono zmienne obrazujące profil gospodarczy, poziom inwestycji, dochodowość przemysłu oraz kondycję finansową przedsiębiorstw (tab. 8). Skoncentrowano się zatem na ekonomicznych aspektach funkcjonowania gospodarki regionalnej. Wybór zmiennych, oprócz wartości merytorycznej, podyktowany był także przydatnością w opisie zjawisk na poziomie regionalnym, a także spełnianiem kryterium porównywalności i mierzalności.

W opisie województw uwzględniono dane dla podregionów, na podstawie których wyznaczono przedziały liczbowe. Tablicę danych symbolicznych zaprezentowano w tab. 9.

**Tabela 8.** Zbiór zmiennych symbolicznych

Skrócona nazwa zmiennej	Definicja zmiennej	Rodzaj zmiennej symbolicznej	Zbiór realizacji zmiennej
Inwestycje	nakłady inwestycyjne w przedsiębiorstwach na 1 mieszkańca (zł)	przedział liczbowy	[729,00; 11 798]
Usługi	udział pracujących w usługach i handlu* w pracujących ogółem (%)	przedział liczbowy	[26,30; 85,61]
Przemysł	produkcja sprzedana przemysłu ogółem na 1 mieszkańca (zł)	przedział liczbowy	[5 052,00; 97 766,00]
Wynagrodzenia	przeciętne miesięczne wynagrodzenie brutto (zł)	przedział liczbowy	[2 746,13; 4 936,36]

\* naprawa pojazdów samochodowych; transport i gospodarka magazynowa; zakwaterowanie i gastronomia; handel; informacja i komunikacja; działalność finansowa i ubezpieczeniowa; obsługa rynku nieruchomości, pozostałe usługi

Źródło: opracowanie na podstawie danych BDL GUS.

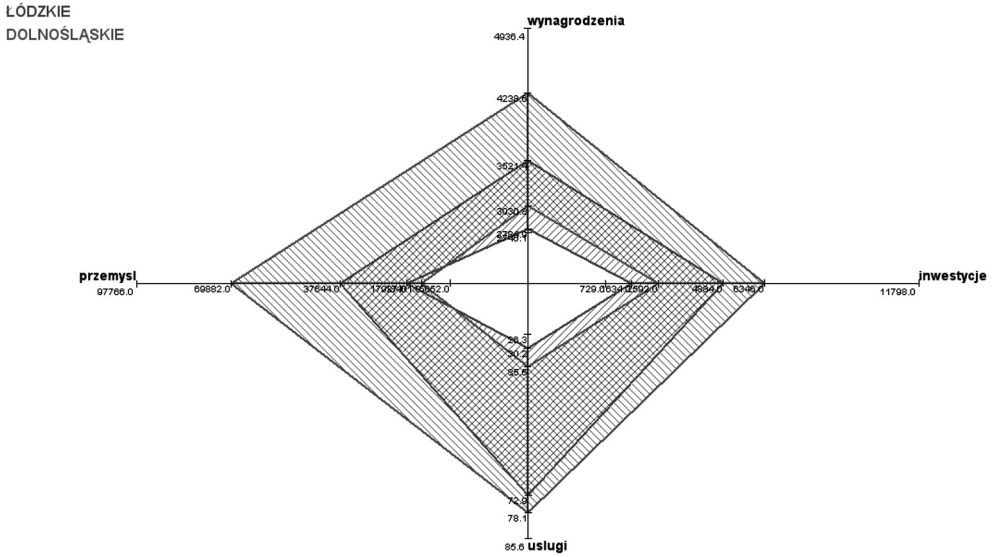
**Tabela 9.** Tablica danych symbolicznych

	inwestycje	usługi	przemysł	wynagrodzenia
ŁÓDZKIE	[ 1634.00 : 4884.00 ]	[ 30.20 : 72.91 ]	[ 17927.00 : 37644.00 ]	[ 2784.94 : 3521.42 ]
MAZOWECKIE	[ 1466.00 : 11798.00 ]	[ 35.18 : 85.61 ]	[ 14238.00 : 97766.00 ]	[ 3125.97 : 4936.36 ]
MAŁOPOLSKIE	[ 1453.00 : 4629.00 ]	[ 34.71 : 75.93 ]	[ 9319.00 : 35278.00 ]	[ 2801.48 : 3722.48 ]
ŚLĄSKIE	[ 1632.00 : 4629.00 ]	[ 37.27 : 66.66 ]	[ 16058.00 : 66706.00 ]	[ 3021.20 : 4551.91 ]
LUBELSKIE	[ 729.00 : 3665.00 ]	[ 27.29 : 52.24 ]	[ 6054.00 : 18136.00 ]	[ 2923.32 : 3551.29 ]
PODKARPACKIE	[ 1565.00 : 3261.00 ]	[ 33.12 : 44.04 ]	[ 5052.00 : 27488.00 ]	[ 2796.37 : 3220.79 ]
PODLASKIE	[ 2016.00 : 3204.00 ]	[ 30.82 : 56.92 ]	[ 9883.00 : 19607.00 ]	[ 3062.15 : 3255.08 ]
ŚWIĘTOKRZYSKIE	[ 2410.00 : 2799.00 ]	[ 26.30 : 46.75 ]	[ 17346.00 : 17750.00 ]	[ 3134.45 : 3145.71 ]
LUBUSKIE	[ 2399.00 : 4021.00 ]	[ 50.66 : 52.35 ]	[ 19434.00 : 30810.00 ]	[ 3065.04 : 3079.78 ]
WIĘLKOPOLSKIE	[ 2021.00 : 6225.00 ]	[ 34.48 : 76.89 ]	[ 20692.00 : 59866.00 ]	[ 2746.13 : 3987.13 ]
ZACHODNIOPOMORSKI	[ 1676.00 : 3630.00 ]	[ 46.23 : 79.43 ]	[ 9351.00 : 32236.00 ]	[ 2941.43 : 3761.89 ]
DOLNOŚLĄSKIE	[ 2592.00 : 6346.00 ]	[ 35.52 : 78.11 ]	[ 13401.00 : 69882.00 ]	[ 3030.82 : 4238.64 ]
OPOLSKIE	[ 1541.00 : 3581.00 ]	[ 43.80 : 48.71 ]	[ 8581.00 : 25531.00 ]	[ 2927.49 : 3385.41 ]
KUJAWSKO-POMORSKI	[ 2135.00 : 3885.00 ]	[ 37.56 : 60.80 ]	[ 18390.00 : 26114.00 ]	[ 2829.77 : 3279.74 ]
POMORSKIE	[ 1881.00 : 5712.00 ]	[ 43.21 : 76.04 ]	[ 11909.00 : 60078.00 ]	[ 2924.93 : 4189.86 ]
WARMIŃSKO-MAZURSKI	[ 1733.00 : 2528.00 ]	[ 43.37 : 54.29 ]	[ 9922.00 : 21415.00 ]	[ 2806.10 : 3236.59 ]

Źródło: opracowanie na podstawie danych BDL GUS.

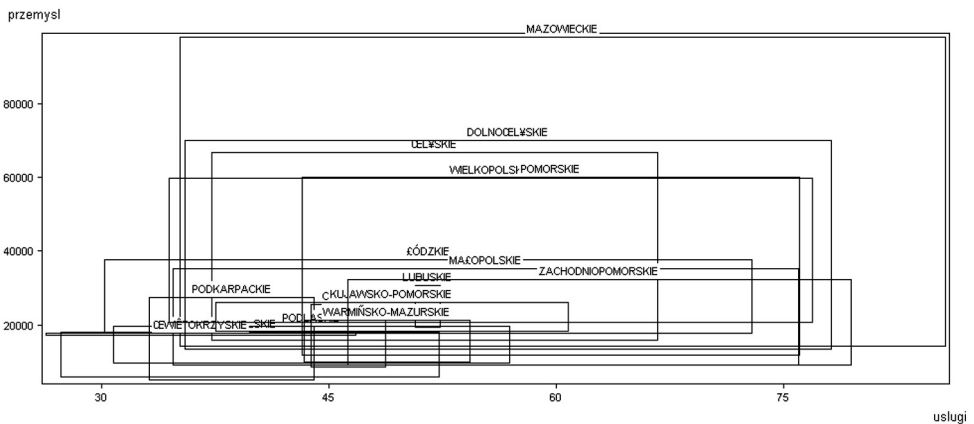
Porównanie wartości zmiennych symbolicznych uzyskanych przez województwa łódzkie (mniejszy wielokąt) i dolnośląskie (większy wielokąt) zaprezentowano za pomocą wykresu rozwiązdy na rys. 3. Województwo dolnośląskie (obejmuje podregiony jeleniogórski, legnicko-głogowski, wałbrzyski, wrocławski i m. Wrocław) jest znacznie bardziej zróżnicowane wewnętrznie pod względem sytuacji go-

spodarczej niż województwo łódzkie (obejmujące podregiony łódzki, piotrkowski, sieradzki, skierniewicki i m. Łódź).



Rys. 3. Wartości zmiennych symbolicznych dla województwa łódzkiego i dolnośląskiego w 2010 r.

Źródło: opracowanie własne na podstawie danych z tab. 10, w pakiecie symbolicDA (Dudek, Pełka, Wilk) programu R.



Rys. 4. Wartości wybranych zmiennych symbolicznych dla województw w 2010 r.


Źródło: opracowanie własne na podstawie danych z tab. 10, w pakiecie symbolicDA (Dudek, Pełka, Wilk) programu R

Rozpatrując dochodowość przemysłu oraz profil gospodarczy, można zauważyć, że największe dysproporcje rozwojowe występują w województwie mazowieckim. Słabo zróżnicowane wewnętrznie są natomiast województwa świętokrzyskie, podkarpackie i warmińsko-mazurskie.

Na podstawie danych z tabeli 10 dokonano pomiaru odległości obiektów (województw) z wykorzystaniem miary odległości Ichino-Yaguchiego (wzory (1) i (2)). Miara przyjmuje wartości w przedziale  $[0, 1]$ , gdzie wartość 0 oznacza bardzo wysokie podobieństwo, natomiast wartość 1 – bardzo duże rozbieżności pod względem sytuacji gospodarczej. W tabeli 10 przedstawiono uzyskane wyniki oraz zaznaczono 5 par województw o najbardziej zbliżonej i rozbieżnej sytuacji gospodarczej.

**Tabela 10.** Macierz odległości obiektów symbolicznych\*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	0,75														
3	0,00	0,71													
4	0,35	0,46	0,32												
5	0,25	0,96	0,28	0,48											
6	0,30	0,98	0,30	0,56	0,10										
7	0,21	0,95	0,26	0,49	0,06	0,08									
8	0,34	1,00	0,45	0,59	0,14	0,11	0,06								
9	0,41	0,96	0,43	0,52	0,27	0,19	0,20	0,21							
10	0,15	0,53	0,18	0,20	0,50	0,50	0,47	0,58	0,54						
11	0,18	0,75	0,06	0,35	0,43	0,46	0,36	0,56	0,32	0,27					
12	0,32	0,40	0,28	0,11	0,57	0,61	0,50	0,64	0,58	0,08	0,31				
13	0,35	0,96	0,32	0,48	0,12	0,07	0,13	0,17	0,09	0,51	0,27	0,56			
14	0,14	0,91	0,17	0,44	0,17	0,14	0,05	0,21	0,15	0,32	0,25	0,46	0,10		
15	0,25	0,49	0,17	0,12	0,52	0,56	0,48	0,65	0,48	0,04	0,17	0,03	0,39	0,38	
16	0,31	0,99	0,29	0,53	0,15	0,12	0,07	0,19	0,12	0,50	0,27	0,59	0,00	0,07	0,43

 duża odległość (małe podobieństwo)

 mała odległość (duże podobieństwo)

\* Wartości zaokrąglono do dwóch miejsc po przecinku. Objasnienia: 1 – łódzkie, 2 – mazowieckie, 3 – małopolskie, 4 – śląskie, 5 – lubelskie, 6 – podkarpackie, 7 – podlaskie, 8 – świętokrzyskie, 9 – lubuskie, 10 – wielkopolskie, 11 – zachodniopomorskie, 12 – dolnośląskie, 13 – opolskie, 14 – kujawsko-pomorskie, 15 – pomorskie, 16 – warmińsko-mazurskie.

Źródło: opracowanie na podstawie danych BDL GUS w pakiecie symbolicDA (Dudek, Pełka, Wilk) programu R.

Analiza wykonana na podstawie danych symbolicznych pozwoliła stwierdzić, że w Polsce występują dysproporcje regionalne, ale należy zauważyć, że większość par województw uzyskała wartości miary niższe niż 0,6. Zatem poziom zróżnicowania regionalnego można określić jako umiarkowany. Jeśli za region wzorcowy uznać



województwo mazowieckie, to najtrudniejsza sytuacja gospodarcza cechuje województwa świętokrzyskie, warmińsko-mazurskie, podkarpackie, lubelskie, lubuskie, opolskie, podlaskie i kujawsko-pomorskie.

## 6. Podsumowanie

Metodologia analizy danych symbolicznych znajduje zastosowanie w przestrzennej analizie sytuacji gospodarczej. Dane symboliczne umożliwiają uwzględnienie struktury rozpatrywanego zjawiska w obrębie porównywanych regionów, szczególnie w sytuacji gdy w regionach występują znaczne dysproporcje wewnętrzne.

Metodyka wypracowana w ramach analizy danych symbolicznych pozwala prowadzić zaawansowane analizy statystyczne na podstawie danych symbolicznych. Macierz odległości, wyznaczona na podstawie danych symbolicznych, może służyć w statystyce i ekonometrii przestrzennej w badaniu podobieństwa regionów oraz zależności przestrzennych.

## Literatura

- Bal-Domańska B., Wilk J., *Gospodarcze aspekty zrównoważonego rozwoju województw – wielowymiarowa analiza porównawcza*, „Przegląd Statystyczny” 2011, nr 3-4, tom 58, s. 300-322.
- Billard L., Diday E., *Symbolic Data Analysis. Conceptual Statistics and Data Mining*, Wiley, Chichester 2006.
- Bock H.H., Diday E. (red.), *Analysis of Symbolic Data. Exploratory Methods for Extracting Statistical Information from Complex Data*, Springer-Verlag, Berlin, Heidelberg 2000.
- Diday E., Noirhomme-Fraiture M. (red.), *Symbolic Data Analysis and the Sodas Software*, John Wiley & Sons, Chichester 2008.
- Everitt B.S., Dunn G., *Applied Multivariate Data Analysis*, Arnold, London 2001.
- Gatnar E., Walesiak M. (red.), *Analiza danych jakościowych i symbolicznych z wykorzystaniem programu R*, C.H. Beck, Warszawa 2011.
- Hair J.F., Black W.C., Babin B.J., Anderson R.E., Tatham R.L., *Multivariate Data Analysis*, Pearson Prentice Hall, New Jersey 2006.
- Ichino M., Yaguchi H., *Generalized Minkowski metrics for mixed feature-type data analysis*, „IEEE Transactions on Systems, Man, and Cybernetics” 1994, vol. 24, no. 4, s. 698-707.
- Malerba D., Esposito F., Giovalle V., Tamma V., *Comparing Dissimilarity Measures for Symbolic Data Analysis*, [w:] *New Techniques and Technologies for Statistics and Exchange of Technology and Know-how*, red. P. Nanopoulos, 2001, s. 473-481.
- Malerba D., Esposito F., Monopoli M., *Comparing Dissimilarity Measures for Probabilistic Symbolic Objects*, [w:] *Data Mining III*, „Series Management Information Systems” vol. 6, red. A. Zanasi, C.A. Brebbia, N.F.F. Ebecken, P. Melli, WIT Press, Southampton 2002, s. 31-40.
- Strahl D. (red.), *Metody oceny rozwoju regionalnego*, Wyd. AE we Wrocławiu, Wrocław 2006.
- Strahl D., *Propozycja miary efektywności innowacyjności w hierarchicznym przekroju regionalnym z wykorzystaniem European Innovation Scoreboard*, [w:] *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 1180, Ekonometria 19, Zastosowania metod ilościowych*, red. J. Dziechciarz, Wrocław 2007, s. 9-18.

- Strahl D., *Strukturalna miara rozwoju obiektów hierarchicznych*, [w:] Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 1100, *Ekonometria* 16, *Zastosowania metod ilościowych*, red. J. Dziechciarz, Wrocław 2006, s. 11-20.
- Wilk J., *Analiza porównawcza oprogramowania komputerowego w klasyfikacji danych symbolicznych*, [w:] Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 242, *Taksonomia* 19, *Klasyfikacja i analiza danych – teoria i zastosowania*, red. K. Jajuga, M. Walesiak, Wrocław 2012a, s. 323-332.
- Wilk J., *Metody analizy danych symbolicznych*, [w:] Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 141, *Ekonometria* 29, *Zastosowania metod ilościowych*, red. J. Dziechciarz, Wrocław 2010, s. 29-38.
- Wilk J., *Miary odległości obiektów opisanych zmiennymi symbolicznymi z wagami*, [w:] Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1126, *Taksonomia* 13, *Klasyfikacja i analiza danych – teoria i zastosowania*, red. K. Jajuga, M. Walesiak, Wrocław 2006b, s. 224-236.
- Wilk J., *Problemy klasyfikacji obiektów symbolicznych. Symboliczne miary odległości*, [w:] *Ilościowe i jakościowe metody badania rynku. Pomiar i jego skuteczność*, red. J. Garczarczyk, ZN AE w Poznaniu nr 71, Poznań 2006a, s. 69-83.
- Wilk J., *Symbolic approach in regional analyses*, „Statistics in Transition – new series” 2012b, vol. 13, nr 3, s. 581-600.
- Wilk J., *Taksonomiczna analiza rynku pracy województw Polski – podejście symboliczne*, [w:] Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 200, *Ekonometria* 34, *Zastosowania metod ilościowych*, red. J. Dziechciarz, Wrocław 2011, s. 26-37.

## SYMBOLIC DATA IN THE ANALYSIS OF REGIONAL DIVERSIFICATION OF ECONOMIC SITUATION

**Summary:** Symbolic data are applied in the spatial analysis of economic situation. They enable to take account of the structure of a phenomenon within compared regions, especially if significant regional disparities occur. The objective of the paper is to discuss the concept of symbolic data and their applications in describing the economic situations of territorial units. The paper shows the problems in measuring the economic situation in comparative research and the concept and construction of symbolic data for describing the situations of territorial units, as well as the methods of their analysis. The article also presents an empirical example of applying symbolic data in the analysis of regional disproportions of the economic situation in Poland.

**Keywords:** symbolic data, economic distance, regional disparities.