

Grażyna Dehnel

Marek Walesiak

Marek Obrębalski

Spójność społeczna województw Polski – analiza porównawcza wyników dla danych klasycznych oraz symbolicznych interwałowych



III Ogólnopolska Konferencja z cyklu

Statystyka publiczna w służbie samorządu terytorialnego

z tematem przewodnim: **Wyzwania innowacyjne**

Wrocław, 26–27 marca 2018 r.

Plan prezentacji

- 1. Charakterystyka badania – cel badania, metody, dane**
- 2. Spójność społeczna – problemy definicyjne**
- 3. Ocena spójności społecznej Polski na tle krajów UE**
- 4. Metodyka badawcza**
- 5. Spójność społeczna województw Polski**
 - 5a. Założenia badania – dobór zmiennych
 - 5b. Wybór optymalnej procedury skalowania wielowymiarowego
 - 5c. Interpretacja wyników
- 6. Wnioski**

1. Charakterystyka badania – cel badania, metody, dane

Cel badania:

Przeprowadzenie analizy porównawczej wyników badania spójności społecznej województw Polski uzyskanych na podstawie **klasycznych danych metrycznych** oraz **danych symbolicznych interwałowych**.

Metody:

skalowanie wielowymiarowe, porządkowanie liniowe, miary położenia, rozproszenia i korelacji

Dane:

Źródło danych – Bank Danych Lokalnych GUS

Rok – 2016 (2014 – 1 zmienna)

Obiekty badania: 16 województw + województwo przeciętne

Zmienne: 26 zmiennych opisujących spójność społeczną

2. Spójność społeczna – problemy definicyjne

200 lat historii – - różne koncepcje

różne: obszary, okresy, poglądy polityczne, metody działania

podejście akademickie

- związane z naukami społecznymi (socjologia i psychologia)
- ✓ procesy integracji społecznej i stabilizacji społecznej
- ✓ wykluczenie społeczne
- wykorzystanie subiektywnych ocen, które mają charakter jakościowy

Pomijane są kwestie definicyjne.

podejście polityczne

- spójność społeczna jako warunek wstępny dobrobytu gospodarczego.
- jego spełnienie zależy jest od przezwyciężenia problemów ekonomicznych i społecznych, wynikających z podziału dochodu, zatrudnienia, kwestii mieszkaniowych, powszechnego dostępu do opieki zdrowotnej i edukacji, udziału w życiu politycznym i obywatelskim
- najpoważniejsze problemy: bezrobocie, ubóstwo, wykluczenie społeczne

2. Spójność społeczna – problemy definicyjne – cd.

pojęcie polityczne

obszar polityczny i społeczny - 1998 (Jenson)

1. przynależność vs. izolacja społeczna
2. integracja vs. wykluczenie
3. partycypacja (uczestnictwo) vs. bierność (nieangażowanie się)
4. uznanie vs. odrzucenie,

obszar polityczny i społeczny + obszar ekonomiczny – 1999 (Bernard)

5. legitymizacja vs. bezprawie / równość vs. nierówność

obszar polityczny i społeczny – obszar ekonomiczny - 2002 (Beauvais i Jenson)

1. wspólnych wartości i kultury obywatelskiej
2. porządku społecznego i kontroli społecznej
3. solidarności społecznej i zmniejszenia dysproporcji majątkowej
4. sieci społecznej i społecznego kapitału
5. miejsca przywiązania i tożsamości

obszar polityczny i społeczno-kulturowy - 2006 (Chan, J., To, H., & Chan)

2010 (Dickes P., Valentova M., Borsenberger M.) 2013 (Dickes P., Valentova M.)

minimalistyczna wersja definicji pomijająca cechy uznawane za czynniki lub determinanty spójności społecznej
tj. równość szans, równość i integracja społeczna, zaufanie instytucjonalne

1. zaufanie instytucjonalne
2. solidarność,
3. udział społeczno-kulturowy
4. udział w życiu politycznym

2016

regionalny wskaźnik postępu społecznego EU-SPI

2. Spójność społeczna – problemy definicyjne – cd.

regionalny wskaźnik postępu społecznego EU-SPI

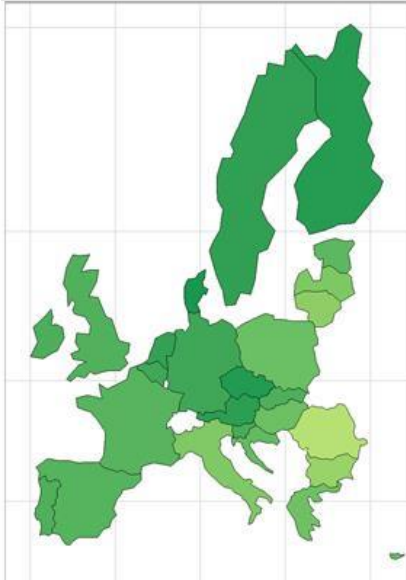
podstawowe potrzeby ludzkie	podstawy dobrostanu	możliwości
odżywianie i podstawowa pomoc medyczna	dostęp do podstawowej wiedzy	prawa osobiste
woda i urządzenia sanitarne	dostęp do informacji i komunikacji	osobista wolność i wybory
schronienie – mieszkanie	zdrowie i rekreacja	tolerancja i włączenie
osobiste bezpieczeństwo	równowaga ekosystemu	prawa osobiste

3. Ocena spójności społecznej Polski na tle krajów UE– wyniki badań

wskaźnik SPI 2016

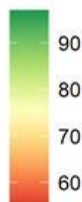


podstawowe potrzeby ludzkie



Wskaźnik postępu społecznego EU-SPI, 2016

Ranking krajów	Kraje	Wskaźnik SPI 2016	Zaspokojenie fundamentalnych potrzeb człowieka	Fundament y dobrobytu	Możliwości awansu społecznego i wolności osobiste
1.	Finlandia	90	96	88	87
2.	Dania	89	97	89	83
3.	Szwecja	89	95	89	82
4.	Holandia	89	95	89	82
5.	W. Brytania	89	93	88	85
6.	Irlandia	88	93	85	85
7.	Austria	87	96	87	77
8.	Niemcy	86	94	87	77
9.	Belgia	86	94	85	79
10.	Hiszpania	86	93	88	77
11.	Francja	85	93	87	75
12.	Słowenia	84	94	83	76
13.	Portugalia	84	93	84	74
14.	Czechy	83	96	83	70
15.	Estonia	83	92	83	73
16.	Włochy	82	89	86	72
17.	Cypr	81	92	82	69
18.	Polska	80	91	80	68
19.	Słowacja	79	93	81	63
20.	Grecja	78	91	83	61
21.	Chorwacja	78	92	80	62
22.	Litwa	77	88	77	66
23.	Węgry	77	91	77	63
24.	Łotwa	76	89	81	59
25.	Rumunia	72	84	75	58
26.	Bulgaria	72	87	75	54



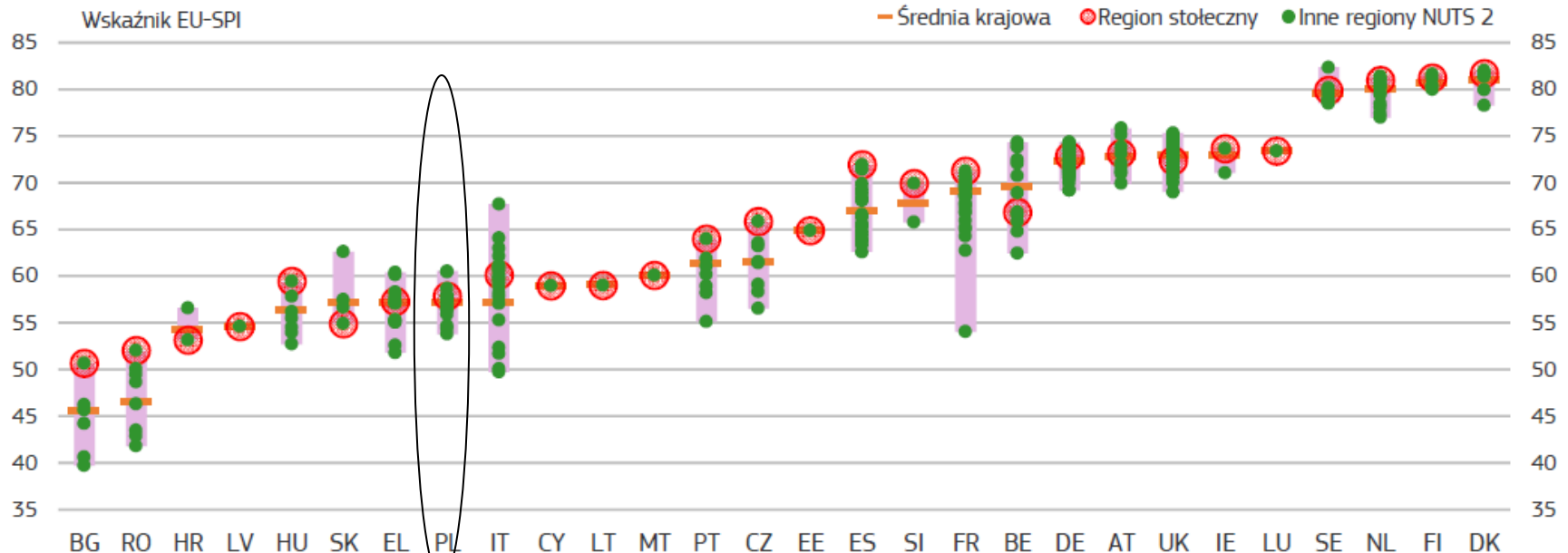
Źródło: „Mój region, moja Europa, nasza przyszłość, Siódmy raport na temat spójności gospodarczej, społecznej i terytorialnej” (2018), Luksemburg: Urząd Publikacji Unii Europejskiej, 2017, <https://publications.europa.eu/en/publication-detail/-/publication/014d75b6-aefa-11e7-837e-01aa75ed71a1/language-pl>

min	72	84	75	54
max	90	97	89	87

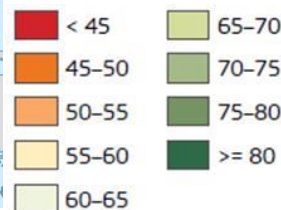
Źródło: Dziembała, M.(2017), Spójność społeczna w krajach i regionach Unii Europejskiej w świetle wskaźnika postępu społecznego, [w:] Nierówności społeczne a wzrost gospodarczy z. 51(3)/2017

3. Ocena spójności społecznej Polski na tle krajów UE

Zmienność EU-SPI w przekroju regionów



Wskaźnik postępu społecznego UE, 2016



Źródło: „Mój region, moja Europa, nasza przyszłość, Siódmy raport na temat spójności gospodarczej, społecznej i terytorialnej” (2018), Luksemburg: Urząd Publikacji Unii Europejskiej, 2017, <https://publications.europa.eu/en/publication-detail/-/publication/014d75b6-aefa-11e7-837e-01aa75ed71a1/language-pl>

4. Metodyka badawcza

W artykule do uporządkowania badanych obiektów ze względu na poziom spójności społecznej zastosowano dwukrokową procedurę badawczą pozwalającą na wizualizację wyników porządkowania liniowego:

- w pierwszym kroku w wyniku zastosowania skalowania wielowymiarowego otrzymuje się wizualizację badanych obiektów w przestrzeni dwuwymiarowej,
- w drugim kroku przeprowadza się porządkowanie liniowe zbioru obiektów.

Przeprowadzono rozszerzoną procedurę badawczą, uwzględniającą osobno:

- **dane metryczne**
- **dane symboliczne interwałowe**

4. Metodyka badawcza cd.

1. Wybrać zjawisko złożone, które nie podlega pomiarowi bezpośredniemu (poziom spójności społecznej).
2. Ustalić zbiór obiektów oraz zbiór zmiennych (metrycznych lub symbolicznych interwałowych) merytorycznie związanych z badanym zjawiskiem złożonym.

Do zbioru obiektów dodaje się wzorzec (górny biegun rozwoju) oraz antywzorzec (dolny biegun rozwoju). Wśród zmiennych wyróżnia się zmienne preferencyjne (stymulanty, destymulanty, nominanty).

3. Zgromadzić dane i zestawić je w postaci:
 - a. Macierzy danych $\mathbf{X} = [x_{ij}]_{n \times m}$ dla zmiennych metrycznych,
 - b. Tablicy danych $\mathbf{X} = [x_{ij}]_{n \times m}$ ($x_{ij} = [x_{ij}^l, x_{ij}^u]$, gdzie $x_{ij}^l \leq x_{ij}^u$) dla zmiennych symbolicznych interwałowych.

Wzorzec (górny biegun) obejmuje najkorzystniejsze wartości zmiennych, antywzorzec (dolny biegun) zaś najmniej korzystne wartości zmiennych preferencyjnych.

W przypadku zmiennych symbolicznych interwałowych współrzędne wyznacza się osobno dla dolnej i górnej wartości przedziału.

4. Metodyka badawcza cd.

4. Przeprowadzić normalizację wartości zmiennych oraz zestawić dane w postaci znormalizowanej:
 - a. Macierzy danych $\mathbf{Z} = [z_{ij}]_{n \times m}$ dla zmiennych metrycznych
 - b. Tablicy danych $\mathbf{Z} = [z_{ij}]_{n \times m}$ ($z_{ij} = [z_{ij}^l, z_{ij}^u]$, gdzie $z_{ij}^l \leq z_{ij}^u$, z_{ij} – znormalizowana obserwacja) dla zmiennych symbolicznych interwałowych.

Dane symboliczne interwałowe wymagają specjalnego podejścia do problemu normalizacji wartości zmiennych. Dolną i górną granicę interwału j -tej zmiennej dla n obiektów łączy się w jeden wektor zawierający $2n$ obserwacji. Takie podejście umożliwia zastosowanie metod normalizacji właściwych dla klasycznych danych metrycznych.

Normalizację zmiennych metrycznych przeprowadzono z wykorzystaniem funkcji **data.Normalization**, a zmiennych symbolicznych interwałowych z wykorzystaniem funkcji **interval_normalization** pakietu **clusterSim**.

4. Metodyka badawcza cd.

W badaniu zastosowano następujące metody normalizacji wartości zmiennych

$$z_{ij} = b_j x_{ij} + a_j = \frac{x_{ij} - A_j}{B_j} = \frac{1}{B_j} x_{ij} - \frac{A_j}{B_j} \quad (b_j > 0)$$

Typ	Nazwa formuły	Parametr	
		B_j	A_j
n1	Standaryzacja	s_j	$\bar{x}_j$
n2	Standaryzacja pozycyjna	mad_j	med_j
n3	Unitaryzacja	r_j	$\bar{x}_j$
n5	Normalizacja w przedziale $[-1; 1]$	$\max_i x_{ij} - \bar{x}_j $	$\bar{x}_j$
n5a	Normalizacja pozycyjna w przedziale $[-1; 1]$	$\max_i x_{ij} - med_j $	med_j
n12a	Normalizacja pozycyjna	$\sqrt{\sum_{i=1}^n (x_{ij} - med_j)^2}$	med_j

4. Metodyka badawcza cd.

5. Wybrać miarę odległości dla **danych metrycznych** (miejska, Euklidesa, Czebyszewa, kwadrat Euklidesa, GDM1), obliczyć odległości i zestawić w macierz odległości $\delta = [\delta_{ik}(\mathbf{Z})]_{n \times n}$ ($i, k = 1, \dots, n$).

Dla **danych symbolicznych interwałowych** wybrać miarę odległości (Ichino-Yaguchi, euklidesowa Ichino-Yaguchi, Hausdorffa, euklidesowa Hausdorffa), obliczyć odległości i zestawić w macierz $\delta = [\delta_{ik}(\mathbf{Z})]_{n \times n}$.

Symbol	Nazwa	Miara odległości $\delta_{ik}(\mathbf{Z})$
U_2_q1	Ichino-Yaguchi 1 $q = 1, \gamma = 0,5$	$\sum_{j=1}^m \varphi(z_{ij}, z_{kj})$
U_2_q2	Euklidesowa Ichino-Yaguchi $q = 2, \gamma = 0,5$	$\sqrt{\sum_{j=1}^m \varphi(z_{ij}, z_{kj})^2}$
H_q1	Hausdorffa $q = 1$	$\sum_{j=1}^m [\max(z_{ij}^l - z_{kj}^l , z_{ij}^u - z_{kj}^u)]$
H_q2	Euklidesowa Hausdroffa $q = 2$	$\left\{ \sum_{j=1}^m [\max(z_{ij}^l - z_{kj}^l , z_{ij}^u - z_{kj}^u)]^2 \right\}^{1/2}$

$z_{ij} = [z_{ij}^l, z_{ij}^u]$; $\varphi(z_{ij}, z_{kj}) = |z_{ij} \oplus z_{kj}| - |z_{ij} \otimes z_{kj}| + \gamma(2 \cdot |z_{ij} \otimes z_{kj}| - |z_{ij}| - |z_{kj}|)$; $| \quad |$ – długość interwału; $z_{ij} \oplus z_{kj} = z_{ij} \cup z_{kj}$; $z_{ij} \otimes z_{kj} = z_{ij} \cap z_{kj}$

6. Przeprowadzić skalowanie wielowymiarowe (MDS): $f: \delta_{ik}(\mathbf{Z}) \rightarrow d_{ik}(\mathbf{V})$ dla wszystkich par (i, k) , gdzie f oznacza odwzorowanie odległości z przestrzeni m -wymiarowej $\delta_{ik}(\mathbf{Z})$ w odpowiadające odległości $d_{ik}(\mathbf{V})$ w przestrzeni q -wymiarowej ($q < m$).

Ze względu na możliwość graficznej prezentacji wyników $q = 2$. Odległości $d_{ik}(\mathbf{V})$ są nieznane. Iteracyjny schemat postępowania w algorytmie `smacof`, pozwalający na znalezienie konfiguracji $\mathbf{V}$ (przy ustalonej liczbie wymiarów q) oraz obliczenie macierzy odległości $d_{ik}(\mathbf{V})$, przedstawiono w pracy (Borg, Groenen 2005, s. 204-205).

W artykule zastosowano rozwiązanie pozwalające na wybór optymalnej procedury skalowania wielowymiarowego, ze względu na zastosowane metody normalizacji wartości zmiennych, miary odległości i modele skalowania, zgodnie z procedurami (dla danych metrycznych oraz symbolicznych interwałowych) dostępnymi w pakiecie `mdsOpt` (Walesiak, Dudek 2017b; 2017c) programu R, które wykorzystują funkcję `smacofSym` pakietu `smacof` (Mair i in., 2017).

4. Metodyka badawcza cd.

W analizie uwzględniono następujące modele skalowania wielowymiarowego (MDS):

a. przekształcenie ilorazowe (ratio MDS):

$$\hat{d}_{ik} = b \cdot \delta_{ik}$$

b. przekształcenie interwałowe (interval MDS):

$$\hat{d}_{ik} = a + b \cdot \delta_{ik}$$

c. przekształcenie wielomianowe (mspline MDS):

$$\hat{d}_{ik} = a + b \cdot \delta_{ik} + c \cdot \delta_{ik}^2 - \text{wielomian 2 stopnia}$$

$$\hat{d}_{ik} = a + b \cdot \delta_{ik} + c \cdot \delta_{ik}^2 + d \cdot \delta_{ik}^3 - \text{wielomian 3 stopnia}$$

7. Ostatecznie w wyniku zastosowania skalowania wielowymiarowego otrzymuje się macierz danych w przestrzeni dwuwymiarowej $V = [v_{ij}]_{n \times q}$ ($q = 2$).

W zależności od położenia wzorca i antywzorca w dwuwymiarowej przestrzeni skalowania $V = [v_{ij}]_{n \times 2}$ wymagany jest obrót układu współrzędnych o kąt φ zgodnie ze wzorem:

$$[v'_{ij}]_{n \times 2} = [v_{ij}]_{n \times 2} \times D$$

gdzie: $[v'_{ij}]_{n \times 2}$ – macierz danych w dwuwymiarowej przestrzeni skalowania po obrocie układu współrzędnych o kąt φ ,

$$D = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \text{ – macierz obrotu.}$$

Obrót ten nie zmienia rozmieszczenia obiektów względem siebie. Pozwala na umieszczenie osi zbioru, łączącej antywzorzec i wzorzec, wzdłuż głównej przekątnej, co daje korzystniejszą wizualizację wyników.

4. Procedura badawcza cd.

8. Przedstawić graficznie oraz zinterpretować wyniki w przestrzeni dwuwymiarowej (wyniki skalowania wielowymiarowego).

Na rysunku w przestrzeni dwuwymiarowej łączy się linią prostą punkty oznaczające antywzorzec i wzorzec w tzw. oś zbioru. Wyznacza się od punktu wzorca izokwanty rozwoju (krzywe jednakowego rozwoju). Obiekty znajdujące się pomiędzy izokwantami prezentują zbliżony poziom rozwoju. Ten sam poziom rozwoju mogą osiągnąć obiekty znajdujące się w różnych punktach na tej samej izokwancie rozwoju (z uwagi na inną konfigurację wartości zmiennych).

9. Uporządkować obiekty na podstawie wartości miary agregatowej d_i bazującej na odległości Euklidesa od obiektu wzorca (Hellwig, 1981):

$$d_i = 1 - \sqrt{\sum_{j=1}^2 (v_{ij} - v_{+j})^2} / \sqrt{\sum_{j=1}^2 (v_{+j} - v_{-j})^2}$$

v_{ij} – j -ta współrzędna dla i -tego obiektu w dwuwymiarowej przestrzeni MDS,

v_{+j} (v_{-j}) – j -ta współrzędna dla obiektu wzorca (antywzorca) w dwuwymiarowej przestrzeni MDS.

Wartości miary agregatowej d_i należą do przedziału $[0; 1]$. Im wyższa jest wartość d_i , tym wyższy jest poziom spójności społecznej badanych obiektów. Porządkuje się obiekty badania według malejących wartości miary agregatowej d_i .

5. Spójność społeczna województw Polski

5a. Założenia badania – dobór zmiennych

Dobór zmiennych wg 3 wymiarów *spójności społecznej*

Podstawowe potrzeby ludzkie

(7 zm.)

Podstawy dobrostanu

(11 zm.)

Możliwości

(8 zm.)

x1 - Przeciętne miesięczne wynagrodzenie brutto (w zł)	x12 - Korzystający z oczyszczalni ścieków (w % ogółu ludności)	x2 - Osoby w gospodarstwach domowych (poniżej kryterium dochodowego) korzystające z pomocy społecznej na 1.000 ludności
x5 - Stopa bezrobocia ogółem w %	x13 - Odsetek ogółu mieszkań wyposażonych w centralne ogrzewanie	x3 - Wskaźnik obciążenia demograficznego (ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym)
x9 - Przeciętna powierzchnia użytkowa mieszkania na mieszkańca w m2	x16 - Dzieci objęte opieką w żłobkach na 1.000 dzieci w wieku do 3 lat	x4 - Udział kobiet wśród pracujących w % (o wartości nominalnej 50%)
x10 - Przeciętna liczba osób na izbę	x17 - Dzieci w placówkach wychowania przedszkolnego na 1.000 dzieci w wieku 3,5 lat	x6 - Udział młodzieży (w wieku do 25 lat) wśród zarejestrowanych bezrobotnych w %
x11 - Długość sieci kanalizacyjnej w relacji do długości sieci wodociągowej w %	x18 - Uczniowie uczący się obowiązkowo języka angielskiego w szkołach podstawowych i gimnazjalnych dla dzieci i młodzieży (w % ogółu uczniów)	x7 - Udział długotrwale bezrobotnych (powyżej 12 miesięcy) wśród zarejestrowanych bezrobotnych w %
x14 - Lekarze i dentyści na 10 tys. ludności	x19 - Uczniowie w liceach ogólnokształcących dla młodzieży na klasę	x8 - Oferty pracy dla osób niepełnosprawnych na 1.000 zarejestrowanych bezrobotnych niepełnosprawnych
x25 - Przestępstwa stwierdzone (kryminalne przeciwko życiu i zdrowiu przeciwko mieniu) na 10 tys. Ludności	x20 - Ćwiczący w klubach sportowych na 1.000 ludności	x15 - Miejsca w placówkach stacjonarnej pomocy społecznej na 10 tys. ludności
	x21 - Czytelnicy bibliotek publicznych na 1.000 ludności	x26 - Frekwencja wyborcza do rad gmin i rad miast w miastach na prawach powiatu w %
	x22 - Uczestnicy imprez kulturalnych na 1.000 ludności	
	x23 - Powierzchnia zieleni publicznej na 10 tys. ludności (w ha)	

- Bank Danych Lokalnych GUS
- dane dla 2016 roku (2014 – zmienna x26)

5. Spójność społeczna województw Polski

5a. Założenia badania – dobór zmiennych

Zmienną nominantę **x4** zamieniono na stymulantę zgodnie z formułą różnicową (Walesiak, 2016a, s. 19):

$$x_{ij} = -|x_{ij}^N - nom_j|$$

gdzie: x_{ij}^N – wartość j -tej nominanty zaobserwowana w i -tym obiekcie,

nom_j – nominalny poziom j -tej zmiennej

x4 – Udział kobiet wśród pracujących w % (*nominanta o wartości nominalnej 50%*)

5. Spójność społeczna województw Polski

5b. Wybór optymalnej procedury skalowania wielowymiarowego

Dwa podejścia do oceny spójności społecznej województw Polski:

- 1. Podejście klasyczne** – procedurę badawczą z części 4 zastosowano do macierzy danych obejmującej 17 obiektów (16 województw plus województwo przeciętne) opisanych 26 zmiennymi metrycznymi.
- 2. Podejście symboliczne interwałowe**

W celu otrzymania danych symbolicznych interwałowych zastosowano dwustopniowe gromadzenie danych.

Najpierw zgromadzono klasyczne dane metryczne dotyczące spójności społecznej według powiatów Polski (380 powiatów opisanych 26 zmiennymi), a następnie poddano je agregacji do poziomu województw otrzymując dane symboliczne interwałowe.

Dolną granicę interwału dla każdej zmiennej w województwie uzyskano obliczając kwartył pierwszy na podstawie danych z powiatów. Z kolei górną granicę interwału otrzymano obliczając kwartył trzeci.

5. Spójność społeczna województw Polski

5b. Wybór optymalnej procedury skalowania wielowymiarowego

Dla klasycznych danych metrycznych w wyborze optymalnej procedury skalowania uwzględniono:

- **6 metod normalizacji** (n1, n2, n3, n5, n5a, n12a)
- **5 miar odległości** (miejska, Euklidesa, kwadrat Euklidesa, Czebyszewa, GDM1)
- **4 modele skalowania** (przekształcenie ilorazowe, przekształcenie interwałowe, wielomian 2 i 3 stopnia)

otrzymując **120 procedur** skalowania wielowymiarowego.

W wyniku zastosowania pakietu **mdsOpt** programu R wybrano optymalną procedurę skalowania wielowymiarowego obejmującą:

- normalizację pozycyjną (metoda normalizacji n12a),
- model skalowania (interval),
- miara odległości Euklidesa.

Dla danych symbolicznych interwałowych w wyborze optymalnej procedury skalowania uwzględniono:

- **6 metod normalizacji** (n1, n2, n3, n5, n5a, n12a),
- **4 miary odległości** (Ichino-Yaguchi, euklidesowa Ichino-Yaguchi, Hausdorffa, euklidesowa Hausdorffa),
- **4 modele skalowania** (przekształcenie ilorazowe, przekształcenie interwałowe, wielomian 2 i 3 stopnia)

otrzymując **96 procedur** skalowania wielowymiarowego.

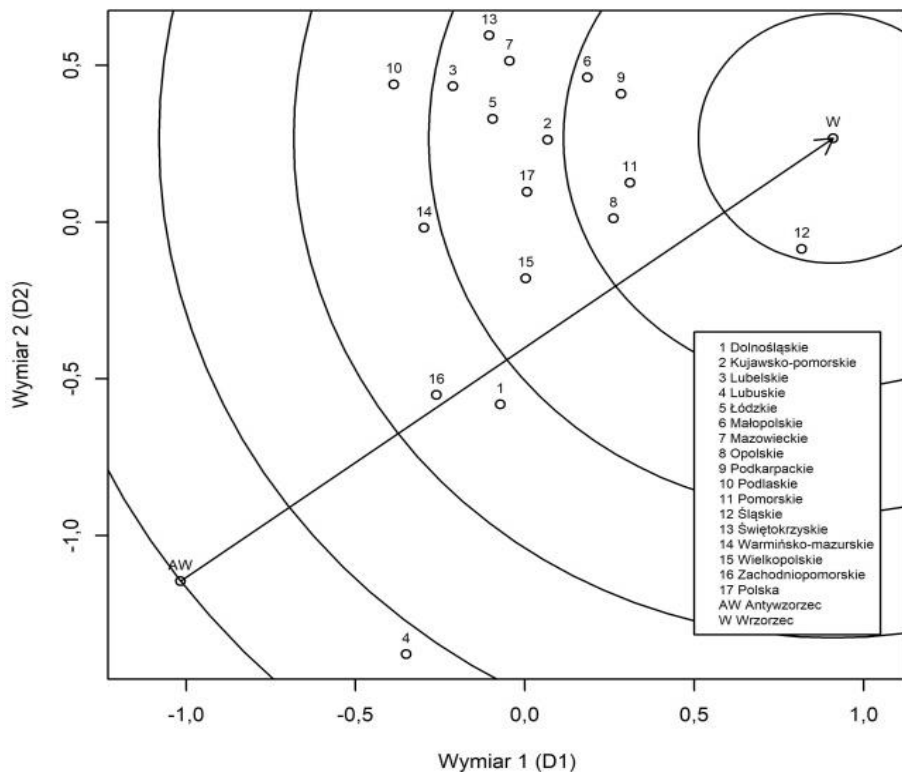
W wyniku zastosowania funkcji **optSmacofSymInterval** pakietu **mdsOpt** programu R wybrano optymalną procedurę skalowania wielowymiarowego. Jest to procedura obejmująca:

- standaryzację pozycyjną (metoda normalizacji n2),
- model skalowania (ratio),
- odległość euklidesowa Hausdorffa.

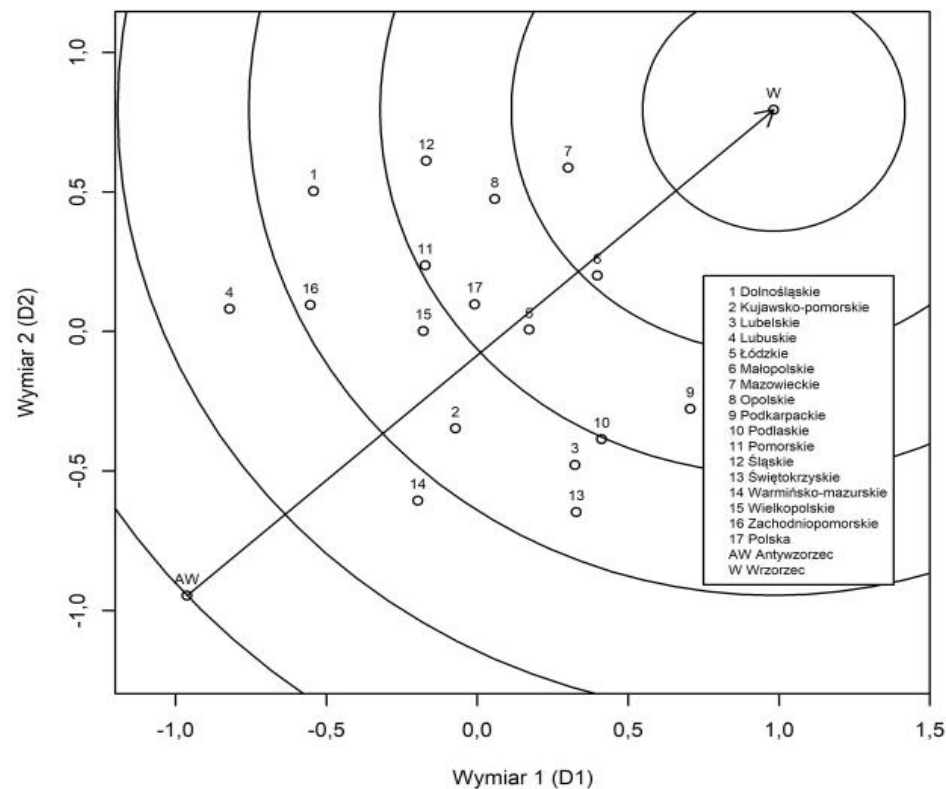
5. Spójność społeczna województw Polski

5c. Interpretacja wyników

dla danych symbolicznych interwałowych



dla danych klasycznych



Zaproponowana modyfikacja metody pozwoliła na przeprowadzenie bardziej dogłębnej oceny wielowymiarowego zjawiska jakim jest spójność społeczna w porównaniu z ujęciem klasycznym.

Województwo nie było oceniane jedynie na podstawie średniej, w jego ocenie istotną rolę odegrało zróżnicowanie powiatów województwa pod względem badanych zmiennych.

Zastosowanie modyfikacji doprowadziło do kluczowych zmian w położeniu obiektów.

5. Spójność społeczna województw Polski

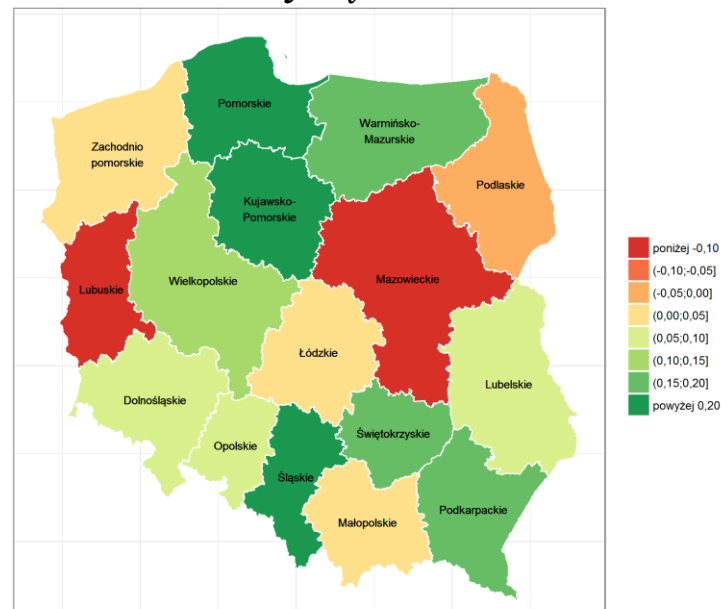
5c. Interpretacja wyników

Uporządkowanie województw Polski, w podejściu klasycznym i symbolicznym interwałowym, ze względu na stopień spójności społecznej w roku 2016 (wartości miary d_i)

Województwo	Nr	Podejście symboliczne interwałowe		Podejście klasyczne		Δd_i
		d_i	Pozycja	d_i	Pozycja	
Śląskie	12	0,8473	1	0,5534	6	0,2939
Pomorskie	11	0,7420	2	0,5086	8	0,2334
Podkarpackie	9	0,7313	3	0,5755	4	0,1557
Opolskie	8	0,7082	4	0,6251	3	0,0831
Małopolskie	6	0,6856	5	0,6806	2	0,0050
Kujawsko-pomorskie	2	0,6472	6	0,4042	13	0,2430
Polska (województwo przeciętne)	17	0,6150	7	0,5356	7	0,0793
Mazowieckie	7	0,5869	8	0,7269	1	-0,1400
Łódzkie	5	0,5784	9	0,5667	5	0,0117
Wielkopolskie	15	0,5766	10	0,4614	10	0,1152
Świętokrzyskie	13	0,5533	11	0,3934	14	0,1600
Lubelskie	3	0,5252	12	0,4507	11	0,0745
Warmińsko-mazurskie	14	0,4806	13	0,2983	16	0,1823
Dolnośląskie	1	0,4565	14	0,4051	12	0,0514
Podlaskie	10	0,4524	15	0,4975	9	-0,0452
Zachodniopomorskie	16	0,4018	16	0,3534	15	0,0484
Lubuskie	4	0,1323	17	0,2570	17	-0,1246
Parametry		Wartość		Wartość		Różnica
Średnia		0,5718	X	0,4878	X	0,0840
Odchylenie standardowe		0,1591	X	0,1252	X	0,0339
Mediana		0,5784		0,4975		0,0809
Medianowe odchylenie bezwzględne		0,1589	X	0,1370	X	0,0219
Współczynnik korelacji liniowej Pearsona		0,6780				
Współczynnik tau Kendalla		0,4853				

Zgodność uporządkowania województw według wartości miary d_i (mierzona współczynnikiem korelacji tau Kendalla $r_{\tau} = 0,4853$) jest znacznie mniejsza niż skorelowanie województw ze względu na wartość miary d_i (mierzone współczynnikiem korelacji liniowej Pearsona $r = 0,6780$).

Zatem porządek województw zmienił się w większym stopniu niż wartości miary d_i .

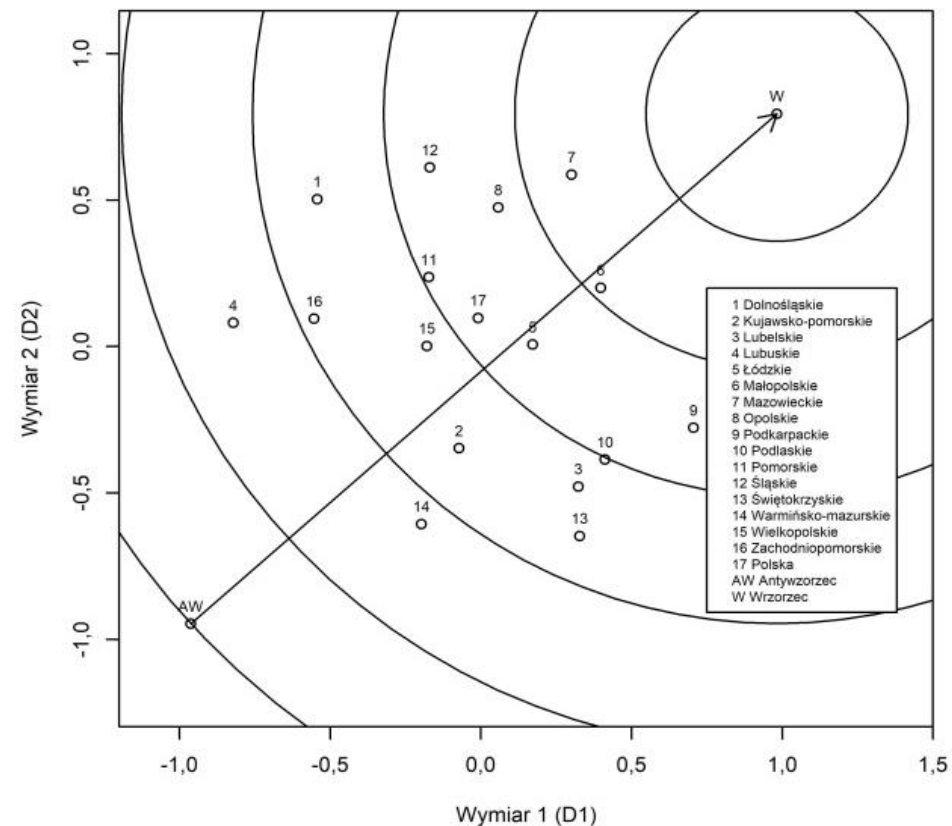
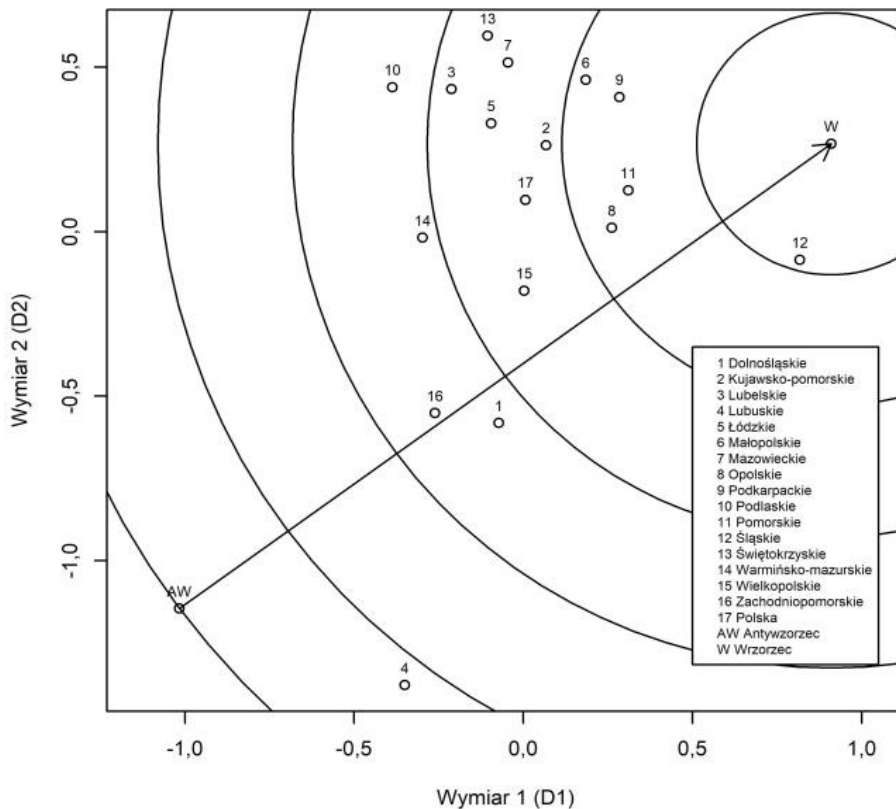


5. Spójność społeczna województw Polski

5c. Interpretacja wyników

dla danych symbolicznych interwałowych

dla danych klasycznych



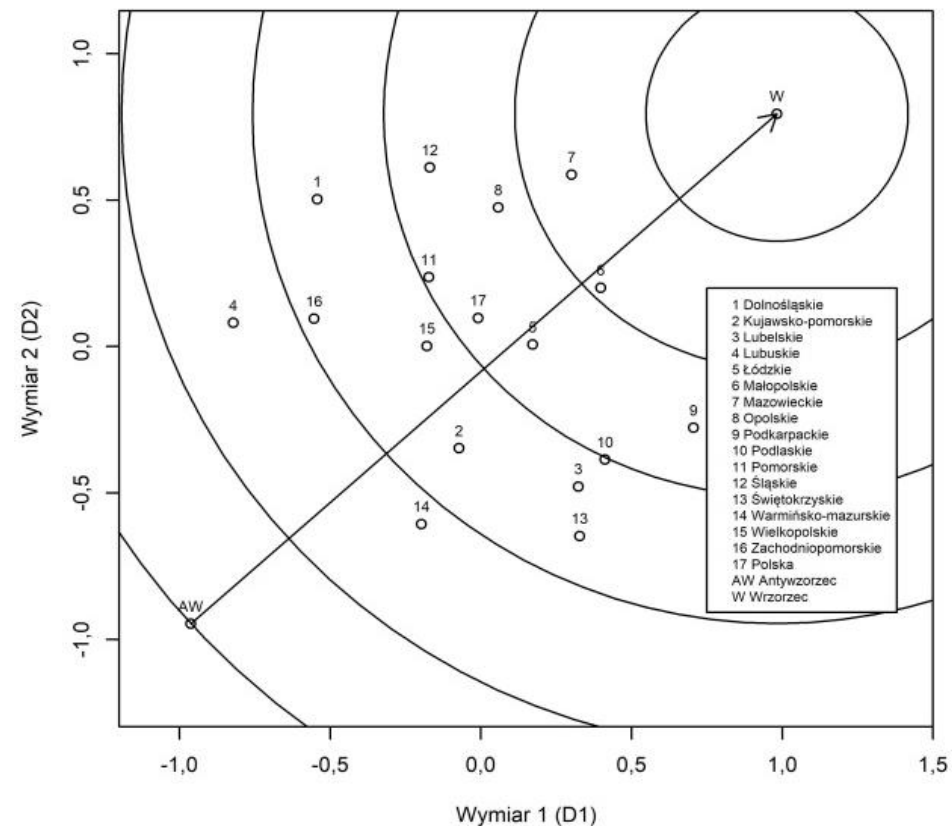
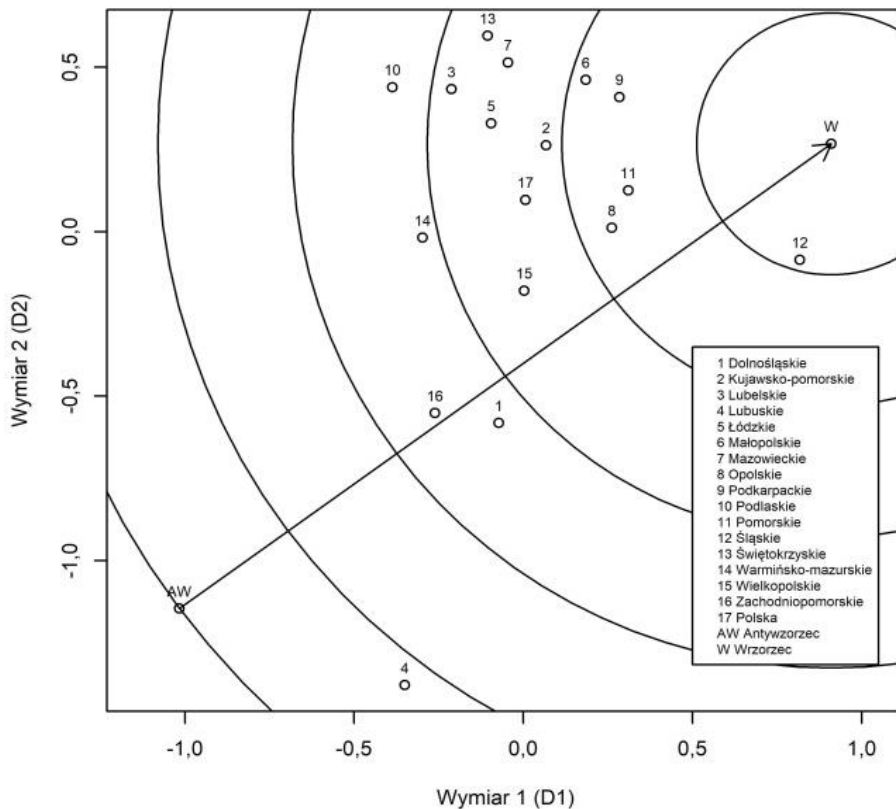
1. Zanotowano zmiany w rozproszeniu obiektów. Zwiększyła się dyspersja województw mierzona odchyleniem standardowym oraz medianowym odchyleniem bezwzględnym. Wyłączenie z analizy województwa lubuskiego (obserwacja odstająca nr 4) spowoduje, że zmienność miary d_i w obu ujęciach będzie kształtowała się na zbliżonym poziomie $S_{d_i} \approx 0,12$.

5. Spójność społeczna województw Polski

5c. Interpretacja wyników

dla danych symbolicznych interwałowych

dla danych klasycznych

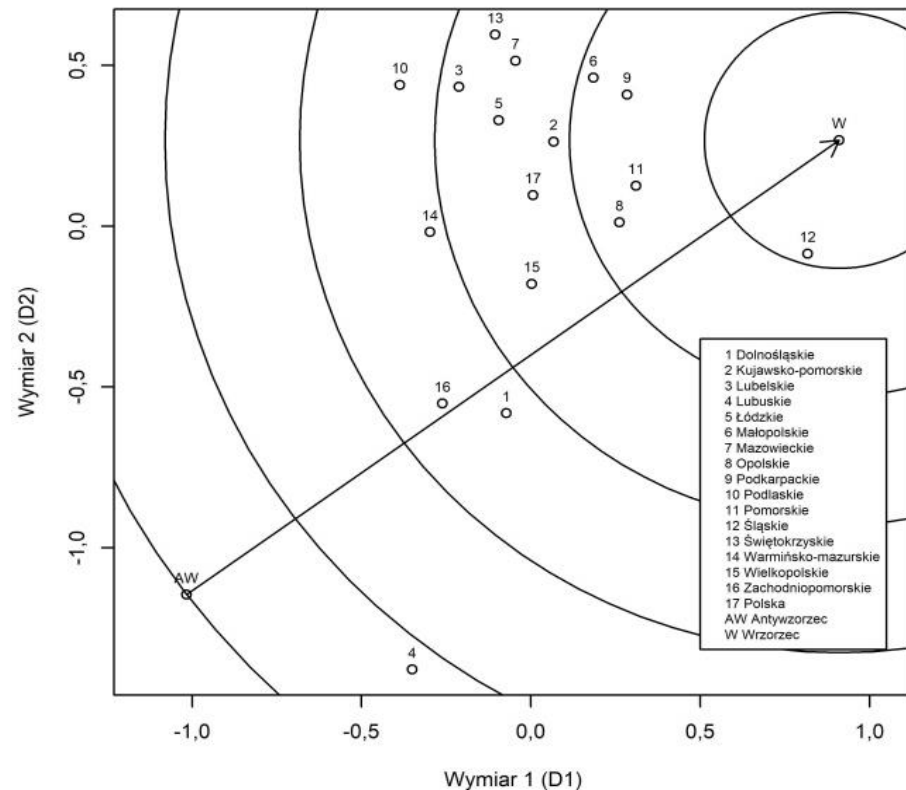


2. Oceniając rozmieszczenie województw nie tylko pod kątem wartości d_i , ale również ich położenia względem osi zbioru można zauważyć, że w przypadku ujęcia symbolicznego interwałowego rozproszenie obiektów jest znacznie mniejsze. Wynika to z pominięcia w analizie wpływu skrajnych wartości badanych zmiennych.

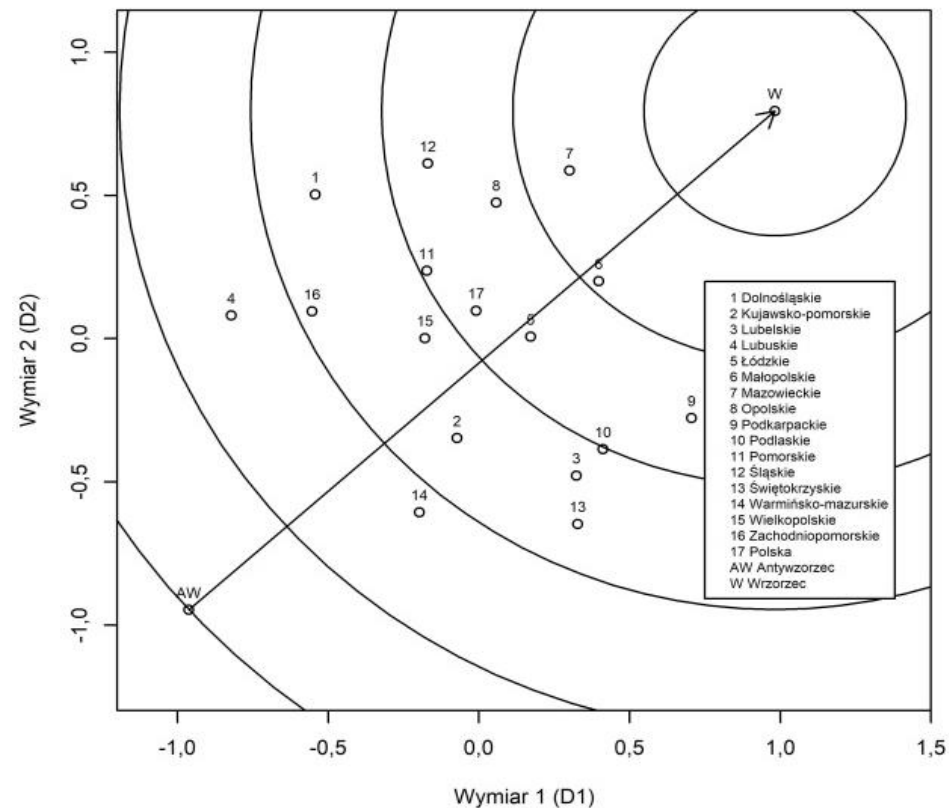
5. Spójność społeczna województw Polski

5c. Interpretacja wyników

dla danych symbolicznych interwałowych



dla danych klasycznych



3. Zdecydowana większość województw (13 województw oraz województwo przeciętne) przesunęła się w kierunku wzorca. Zwiększył się średni poziom spójności społecznej oraz wartość mediany.

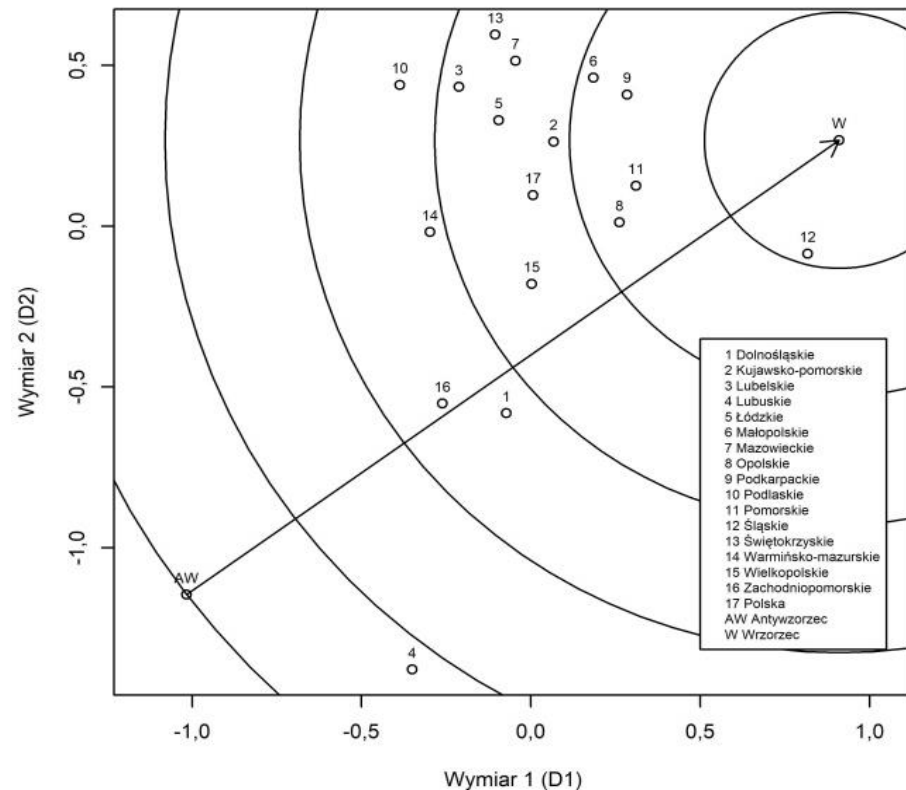
Wśród województw, w których zanotowano największy przyrost miary d_i znalazły się: śląskie (12), pomorskie (11) i kujawsko-pomorskie (2).

Przesunięcie w kierunku antywzorca zanotowano w przypadku trzech województw: mazowieckiego (7), lubuskiego (4) i podlaskiego (10). Oznaczało to, że zastosowanie modyfikacji wpłynęło na obniżenie oceny poziom spójności społecznej.

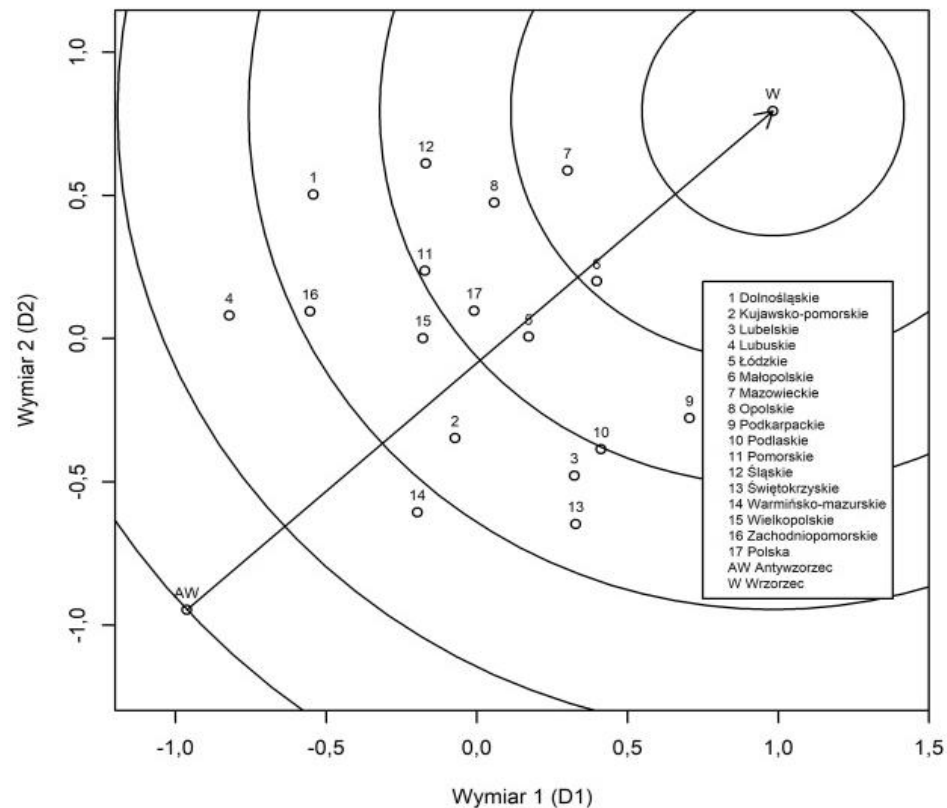
5. Spójność społeczna województw Polski

5c. Interpretacja wyników

dla danych symbolicznych interwałowych



dla danych klasycznych



4. Wykorzystywana w ujęciu klasycznym, jako pojedynczy parametr wartość średnia, jak wiadomo, pozostaje pod silnym wpływem obserwacji odstających. Pod ich silnym wpływem pozostaje ocena całego regionu.

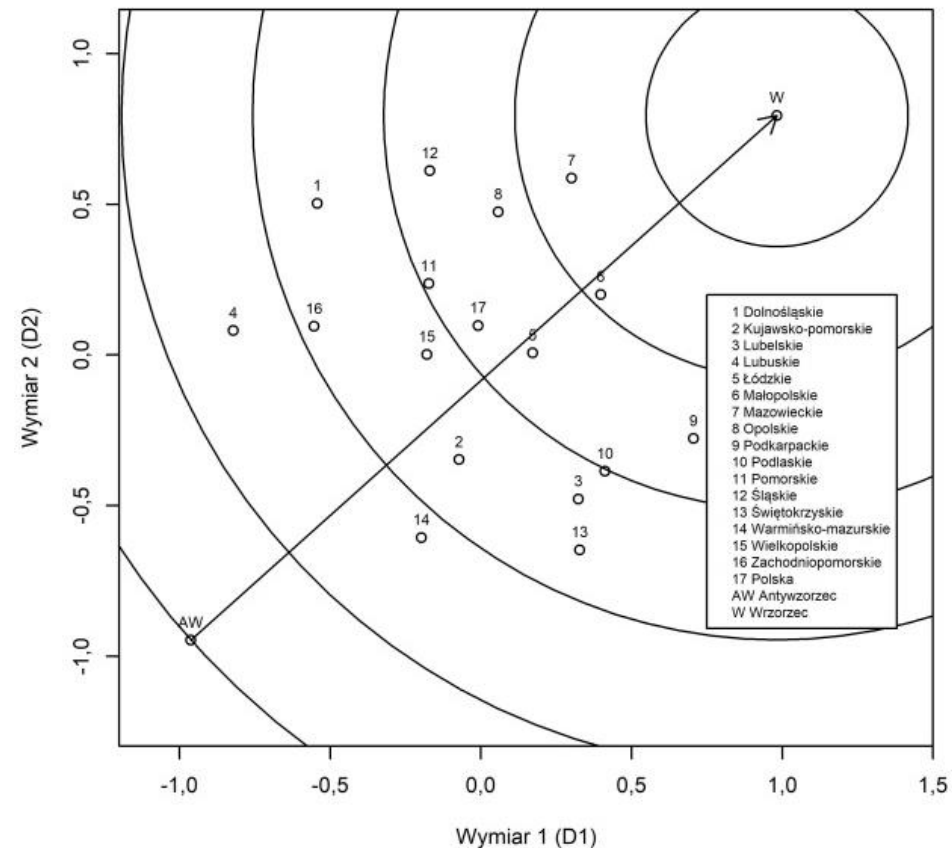
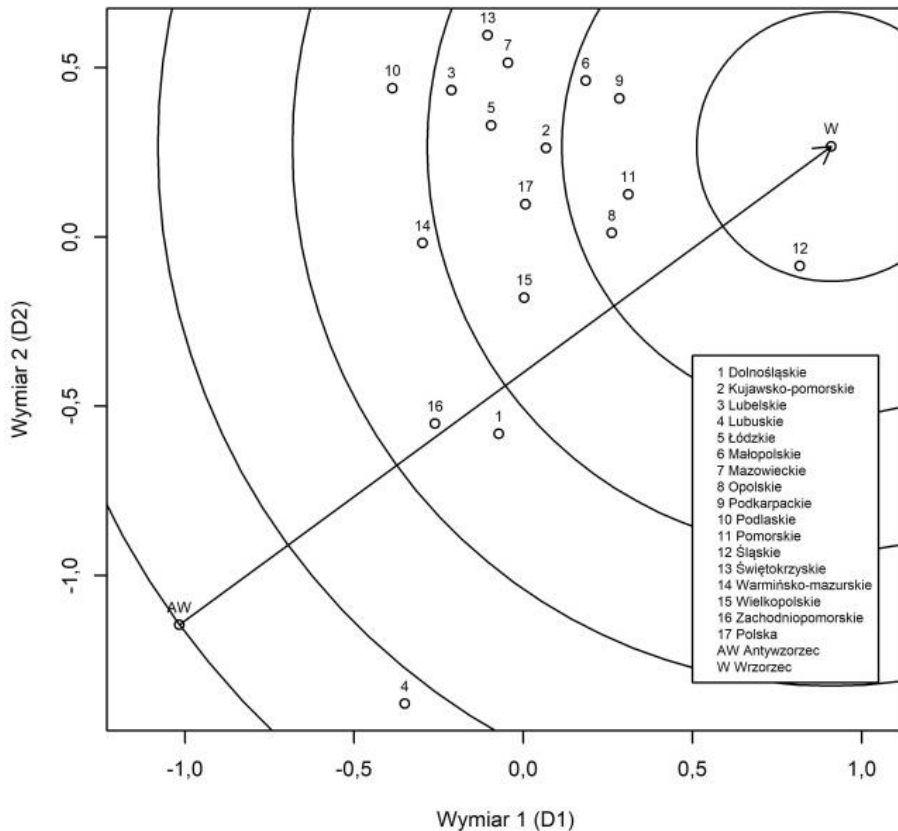
W przeprowadzonym badaniu przykładem jest województwo mazowieckie (7) i oddziaływanie Warszawy oraz województwo lubuskie (4) i oddziaływanie Zielonej Góry. Przejście z ujęcia klasycznego na ujęcie symboliczne interwałowe spowodowało przesunięcie się tych obiektów w kierunku antyzorca. W przypadku województwa mazowieckiego i lubuskiego zanotowano największą tego rodzaju zmianę w położeniu obiektu wyrażoną ujemną wartością przyrostu miary d_i

5. Spójność społeczna województw Polski

5c. Interpretacja wyników

dla danych symbolicznych interwałowych

dla danych klasycznych



5. Wykorzystanie metody bazującej na danych symbolicznych interwałowych spowodowało przesunięcie ośmiu województw powyżej osi zbioru. Poniżej osi zbioru pozostały jedynie trzy województwa.

Zmiana ta wynikała przede wszystkim z uwzględnienia w analizie asymetrii rozkładów badanych zmiennych. Zmienne włączone do badania w przeważającej liczbie województw charakteryzowały się prawostronną asymetrią. Województwa, które znalazły się nad osią zbioru to województwa, które zdominowane były przez takie zmienne. W podejściu klasycznym tego rodzaju szczegółowa ocena uwzględniająca asymetrię rozkładów jest niemożliwa.

6. Wnioski

1. Badanie poziomu spójności społecznej jest złożone i wymaga podejścia wielowymiarowego.
2. Do uporządkowania badanych obiektów ze względu na poziom spójności społecznej zastosowano dwukrokową procedurę badawczą (skalowanie wielowymiarowe + porządkowanie liniowe) pozwalającą na wizualizację wyników w przestrzeni dwuwymiarowej.
3. Dzięki uwzględnieniu dwóch podejść możliwe było ukazanie zmian, jakie w ocenie spójności spowodowało przejście z oceny jednoparametrycznej (dane klasyczne) na przedziałową (dane symboliczne interwałowe).
4. Zaproponowana modyfikacja pozwoliła na ocenę spójności społecznej, której województwo nie było oceniane jedynie na podstawie średnich wartości zmiennych, a w jego ocenie wzięto pod uwagę 50% powiatów (przedział między kwartylem trzecim i pierwszym).
5. Zaproponowane podejście pozwoliło na wyeliminowanie wpływu wartości odstających.

Dziękujemy za
uwagę