

## KLASYFIKACJE OBSZARÓW POLSKI Z PUNKTU WIDZENIA WYBRANYCH CECH SPOŁECZNO-DEMOGRAFICZNYCH LUDNOŚCI

**Anna Majdzińska**

Zakład Demografii i Gerontologii Społecznej, Uniwersytet Łódzki  
e-mail: a\_majdzinska@uni.lodz.pl

**Streszczenie:** W opracowaniu przedstawiono próby klasyfikacji obszarów Polski pod względem wybranych cech społeczno-demograficznych. Rozważono w tym względzie strukturę wykształcenia ludności, w obrębie której szczegółowej analizie poddane zostały struktury wieku i płci ludności legitymującej się wykształceniem wyższym. Badanie przeprowadzono na poziomie województw i powiatów dla lat 2002 i 2011, przy wykorzystaniu wybranych metod analizy skupień.

**Słowa kluczowe:** struktura wykształcenia ludności, struktury wieku i płci ludności z wykształceniem wyższym, metody analizy skupień

### WPROWADZENIE

Celem opracowania były próby klasyfikacji obszarów Polski (województw i powiatów) z punktu widzenia wybranych cech społeczno-demograficznych. Uwaga w tym względzie została zwrócona na strukturę wykształcenia ludności<sup>1</sup>, w obrębie której rozważone zostały struktury wieku i płci ludności legitymującej się wykształceniem wyższym. Wyszczególniono właśnie tę kategorię z tego względu, że w ciągu minionych kilkunastu lat odnotowano w Polsce wyraźny wzrost odsetka ludności legitymującej się wykształceniem wyższym<sup>2</sup>, przy spadku

---

<sup>1</sup> W Polsce struktura wykształcenia ludności badana jest dla ludności w wieku 13+ lat.

<sup>2</sup> Wzrost ten wynikał z chęci kontynuowania nauki, ale także był konsekwencją wprowadzenia w polskich uczelniach wyższych bolońskiego systemu nauczania, mającego dwustopniowy charakter. W rezultacie czego, osoby kończące studia licencjackie, podobnie jak i magisterskie, zaliczane są do frakcji osób posiadających wykształcenie wyższe [zob. Ustawa Prawo o szkolnictwie ..., Dz. U. 2005 nr 164 poz. 1365, art. 2 i 167; MNiSW 2015].

udziałów subpopulacji posiadającej niższe poziomy wykształcenia. W związku z tym interesujące wydaje się być zagadnienie stabilności w czasie struktur demograficznych charakteryzujących subpopulację osób legitymujących się wyższym poziomem wykształcenia.

Badanie przeprowadzono na poziomie województw i powiatów w latach 2002 i 2011, przy wykorzystaniu wybranych metod wielowymiarowej analizy porównawczej (Warda i *k*-średnich). Źródło danych wykorzystanych w analizach stanowiły publikacje Głównego Urzędu Statystycznego dotyczące struktury wykształcenia ludności, pochodzące z wyników spisów powszechnych ludności przeprowadzonych w Polsce w latach 2002 i 2011.

### OPIS STRUKTURY LUDNOŚCI POLSKI WEDŁUG POZIOMU WYKSZTAŁCENIA W LATACH 2002 i 2011

Charakteryzując strukturę wykształcenia ludności można zauważyć, że w Polsce w latach 2002-2011 znacznie wzrósł odsetek ludności z wykształceniem wyższym (o 71,7%). W roku 2011 subpopulacja ta stanowiła 17% ogółu populacji w wieku 13 lat i więcej, podczas gdy w 2002 r. jej udział wyniósł 9,9%, przy czym w obu tych latach omawiana frakcja była znacznie wyższa w miastach (zob. Tabela 1). W obu rozpatrywanych latach odnotowano porównywalny odsetek ludności legitymującej się wykształceniem średnim i policealnym, podobnie jak w przypadku wykształcenia zasadniczego zawodowego, zmniejszyła się natomiast frakcja osób odznaczających się wykształceniem co najwyżej gimnazjalnym<sup>3</sup> i bez wykształcenia (odpowiednio o 22,1% i 61,1%).

Tabela 1. Struktura wykształcenia ludności w wieku 13+ lat w Polsce według płci i według miejsca zamieszkania w latach 2002 i 2011 (w %)

Poziom wykształcenia	Ogółem						Miasta		Wieś	
	Ogółem		Mężczyźni		Kobiety		Ogółem		Ogółem	
	2002	2011	2002	2011	2002	2011	2002	2011	2002	2011
OGÓŁEM	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
wyższe	9,9	17,0	9,3	14,8	10,4	19,0	13,2	21,4	4,2	9,9
średnie i policealne	31,5	31,6	27,6	29,1	35,1	33,8	37,3	35,3	21,5	25,5
zasadnicze zawodowe	23,2	21,7	30,1	27,9	16,9	15,9	20,4	18,6	28,0	26,5
podstawowe ukończone i	29,8	23,2	28,0	22,0	31,4	24,4	23,9	18,0	39,7	31,6
podstawowe nieukończone i bez wykształcenia	3,6	1,4	3,0	1,0	4,3	1,7	2,3	0,9	5,9	2,1
nieustalone	2,0	5,2	2,1	5,2	2,0	5,2	2,8	5,7	0,7	4,3

Źródło: GUS [2013], str. 103-105

<sup>3</sup> W 2002 roku ten poziom wykształcenia nie był badany, gdyż pierwsi absolwenci gimnazjów ukończyli szkołę w czerwcu 2002 roku, czyli później niż moment spisowy, tj. 20 maja 2002 r.

Struktura wykształcenia ludności jest wyraźnie zróżnicowana terytorialnie. W 2011 r. najwyższym odsetkiem osób z wykształceniem wyższym legitymowało się mazowieckie (23,6%), pomorskie (17,6%) i małopolskie (17,2%), najniższym zaś opolskie (13,6%), warmińsko-mazurskie (14,0%) i kujawsko-pomorskie (14,1%), z kolei najwyższym udziałem osób posiadających wykształcenie średnie lub policealne charakteryzowały się województwa mazowieckie (33,7%) i łódzkie (33,1%), najniższym zaś opolskie (28,5%). Największa frakcja osób z wykształceniem zasadniczym zawodowym zamieszkiwała województwa wielkopolskie (26,0%) oraz śląskie i kujawsko-pomorskie (25,2%), najniższa zaś podlaskie (16,0%) i mazowieckie (16,7%). Natomiast najwyższa frakcja ludności o wykształceniu co najwyżej gimnazjalnym występowała w warmińsko-mazurskim (29,7%) i podlaskim (29,5%), najniższa zaś w śląskim (21,3%) oraz mazowieckim i dolnośląskim (22,8%).

W roku 2011, podobnie jak w 2002, najwyższym odsetkiem ludności posiadającej wykształcenie wyższe charakteryzowały się przede wszystkim duże miasta, co tłumaczyć można większą podażą uczelni wyższych i ich lepszą dostępnością w porównaniu do obszarów zlokalizowanych na peryferiach województw. Ludność posiadająca wykształcenie co najmniej średnie koncentrowała się raczej w dużych miastach lub na ich obrzeżach, natomiast osoby z niższymi niż wskazane poziomami wykształcenia skupiały się najczęściej na peryferiach województw.

#### CHARAKTERYSTYKA STRUKTUR DEMOGRAFICZNYCH LUDNOŚCI LEGITYMUJĄCEJ SIĘ WYŻSZYM POZIOMEM WYKSZTAŁCENIA W POLSCE W LATACH 2002 i 2011

Wśród osób legitymujących się wykształceniem wyższym, zarówno w 2011, jak i w 2002 r., przeważała frakcja kobiet (zob. Tabela 2), przy czym w roku 2011 różnica ta uległa zwiększeniu (i wyniosła 16,6 pkt. proc.). W badanym okresie udział kobiet wzrósł o 6,4%, odsetek mężczyzn zaś spadł o 7,7% (zob. Tabela 2).

Tabela 2. Struktura płci ludności z wykształceniem wyższym w Polsce w latach 2002 i 2011 (w %), dynamika zmian w tym okresie oraz różnice pomiędzy udziałami kobiet i mężczyzn w obu latach (w pkt. proc.)

Płeć	2002	2011	2002 = 100
Mężczyźni	45,2	41,7	92,3
Kobiety	54,8	58,3	106,4
Ogółem	100,0	100,0	<b>x</b>

Źródło: GUS [2015], opracowanie własne

W 2011 roku, podobnie jak w 2002, wśród ludności z wyższym poziomem wykształcenia najwyższy udział stanowiły osoby z grupy wieku 25-29 oraz 30-34

lata (zob. Tabela 3). Analiza dynamiki zmian świadczy o tym, że w latach 2002-2011 zwiększeniu uległy udziały osób w wieku 20-39 lat (w największym stopniu wzrósł odsetek osób w wieku 30-34 lata – o 45,6%), zmniejszeniu zaś uległy odsetki osób w wieku 40-59 lat. Zmiany te spowodowały odmłodzenie się frakcji osób legitymujących się wyższym poziomem wykształcenia, w rezultacie czego w latach 2002-2011 mediana wieku tej subpopulacji zmniejszyła się o 4,5 roku.

Przedstawione przemiany w strukturze wieku ludności posiadającej wykształcenie wyższe wynikały przede wszystkim ze znacznego napływu do tej kategorii osób z młodszych grup wieku, ale pośrednio były także następstwem ubytku naturalnego ludności w starszych grupach wieku.

Tabela 3. Struktura wieku ludności legitymującej się wykształceniem wyższym w Polsce w latach 2002 i 2011 (w %), dynamika zmian w tym okresie oraz mediana wieku ludności (w latach)

<b>Grupy wieku</b>	<b>2002</b>	<b>2011</b>	<b>2002 = 100</b>
20-24	6,0	6,8	113,5
25-29	18,7	20,7	111,0
30-34	12,4	18,1	145,6
35-39	10,4	12,4	119,0
40-44	10,8	8,1	75,0
45-49	11,0	6,9	62,0
50-54	9,8	6,9	70,9
55-59	6,4	6,2	96,4
60-64	5,1	5,2	103,3
65 i więcej	9,4	8,7	92,4
<b>Ogółem</b>	<b>100,0</b>	<b>100,0</b>	<b>x</b>
mediana wieku	40,9	36,4	89,0

Źródło: GUS [2015], opracowanie własne

#### KLASYFIKACJE WOJEWÓDZTW I POWIATÓW Z PUNKTU WIDZENIA STRUKTURY WIEKU LUDNOŚCI LEGITYMUJĄCEJ SIĘ WYŻSZYM POZIOMEM WYKSZTAŁCENIA W LATACH 2002 i 2011

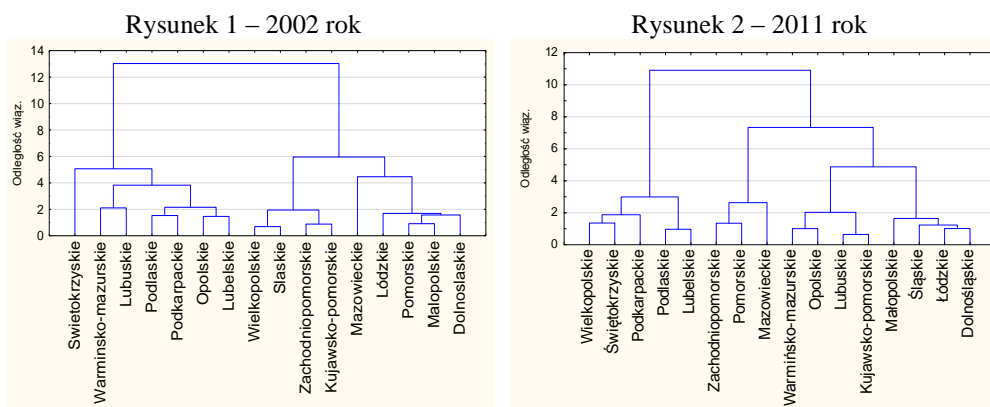
W przypadku struktury wieku ludności legitymującej się wykształceniem wyższym w latach 2002 i 2011 najpierw dokonano prób klasyfikacji województw za pomocą metody Warda z odległością euklidesową jako miarą podobieństwa obiektów, następnie zaś przeprowadzono analizy na poziomie powiatów, stosując w tym przypadku metodę *k*-średnich<sup>4</sup>. W obu przypadkach analizy oparto na procentowych udziałach 5-cioletnich grup wieku ludności.

<sup>4</sup> Metoda Warda oraz metoda *k*-średnich są najczęściej stosowanymi procedurami grupowania. Pierwsza z nich znajduje szczególne zastosowanie w przypadku niezbyt dużego zbioru obiektów, druga zaś w przypadku gdy zbiór ten jest liczny.

Uzyskane za pomocą metody Warda rezultaty hierarchicznego łączenia obiektów dla obu lat zostały przedstawione w postaci drzewka połączeń (zob. Rysunek 1 i 2). Wyniki te różnią się między sobą, co świadczy o zaistniałych w analizowanym okresie zmianach w zakresie podobieństwa województw w badanym względzie. Analiza wykresów przebiegu aglomeracji<sup>5</sup> dla 2002 i 2011 roku wskazała jako najbardziej właściwe podziały na odpowiednio dwa oraz cztery skupienia.

Przedstawione grupowania, ze względu na to, że prowadzone były na dość dużych jednostkach jakimi są województwa, mają raczej pogładowy charakter<sup>6</sup>. W 2011 roku województwami najbardziej podobnymi do siebie z punktu widzenia struktury wieku ludności z wyższym poziomem wykształcenia były m.in.: lubuskie i kujawsko-pomorskie, a także podlaskie i lubelskie (zob. Rysunek 2).

Rysunek 1. i Rysunek 2. Dendrogramy grupowania województw metodą Warda z punktu widzenia struktury wieku ludności z wykształceniem wyższym w latach 2002 i 2011



Źródło: GUS [2015], opracowanie własne

Źródło: GUS [2015], opracowanie własne

<sup>5</sup> Podziału grupy obiektów na skupienia można dokonać w sposób subiektywny, na podstawie wzrokowej analizy dendrytu. Cięcia gałęzi drzewka dokonuje się w miejscu największej różnicy odległości pomiędzy połączeniami [Panek 2009, 120]. Odległość tę można również obliczyć stosując następującą formułę [Panek 2009, 121 za Grabiński]:

$$d_{h-1}^* > \max_h \{d_h / d_{h-1}\} \quad h = 2, 3, \dots, n-1$$

gdzie:  $d_h$  – długość  $h$ -tej gałęzi drzewka,

$d_{h-1}^*$  – wartość krytyczna odległości na wysokości  $h-1$  długości dendrogramu.

Pomocny w tym względzie jest także wykres przebiegu aglomeracji.

W badaniu, opisana formuła wskazała odległości 5,95 oraz 2,99 jako wartości krytyczne.

<sup>6</sup> Uzyskane podziały wydają się być zbyt ogólne i charakteryzują się relatywnie wysokim wewnątrzgrupowym zróżnicowaniem. Schodzenie na niższe poziomy agregacji, tj. dążenie do uzyskania większej liczby skupień, prowadzi do uzyskania kilku jednoelementowych grup. Takie rozwiązanie również nie wydaje się być satysfakcjonujące, gdyż traci się wtedy syntetyczny charakter rezultatów.

Przeprowadzono także grupowanie województw za pomocą metody Warda z punktu widzenia dynamiki zmian w strukturze wieku ludności<sup>7</sup> posiadającej omawianą kategorię wykształcenia w latach 2002-2011. Uzyskane rezultaty pozwoliły wyodrębnić dwie zasadnicze grupy województw, w obrębie których relatywnie wysokim podobieństwem w badanym względzie legitymowały się: warmińsko-mazurskie i podkarpackie, podlaskie i lubelskie, lubuskie i kujawsko-pomorskie, wielkopolskie i łódzkie, a także śląskie i pomorskie. Województwami najbardziej odbiegającymi od pozostałych w zakresie dynamiki omawianych przemian było mazowieckie i świętokrzyskie<sup>8</sup>.

Następnie, przeprowadzono grupowanie powiatów z punktu widzenia struktury wieku ludności legitymującej się wyższym poziomem wykształcenia w latach 2002 i 2011, stosując w tym celu metodę  $k$ -średnich. W przypadku obu tych lat dokonano kolejno podziałów na 4-10 skupień<sup>9</sup>, określonych w oparciu o wyniki aglomeracji metodą Warda. Uzyskane rezultaty zostały poddane ocenie merytorycznej, przy wykorzystaniu wybranych mierników jakości klasyfikacji<sup>10</sup>.

Spśród przeprowadzonych podziałów za najbardziej odpowiedni w obu latach uznano podział na osiem skupień. Wyodrębnione grupy powiatów zostały następnie uporządkowane niemalejąco według mediany wieku ludności. Uzyskane

---

<sup>7</sup> W oparciu o różnice względne w udziałach poszczególnych grup wieku ludności omawianej kategorii wykształcenia, tj. na podstawie formuły:

$(x_1 - x_0) / x_0$ , gdzie:  $x_0$  i  $x_1$  to odsetki danej grupy wieku odpowiednio w latach 2002 i 2011.

<sup>8</sup> Województwa te cechowały się najwyższą wartością średniej z odległości euklidesowych tych jednostek od pozostałych województw.

<sup>9</sup> Liczba mniejsza niż cztery wydaje się być niewystarczająca, natomiast większa niż dziesięć mogłaby okazać się zbyt liczna.

<sup>10</sup> Poprawny podział na skupienia powinien charakteryzować się możliwie niskim zróżnicowaniem wewnątrzgrupowym (czyli wysoką homogenicznością) oraz znacznym zróżnicowaniem międzygrupowym (wysoką heterogenicznością). Ocenę rezultatów (z punktu widzenia tej samej metody i tego samego zjawiska) przeprowadza się za pomocą odpowiednich miar podobieństwa wewnątrz- i międzygrupowego (cząstkowych i sumarycznych) oraz mierników syntetycznych [zob. Grabiński i in. 1989, 150-158; Młodak 2006, 78-81; Nowak 1990, 190-194; Panek 2009, 161-167]. W niniejszym opracowaniu jako miarę homogeniczności wykorzystano średnią odległość obiektów danej grupy od jej środka ciężkości, natomiast jako miarę heterogeniczności zastosowano średnią odległość między środkiem ciężkości danego skupienia a centrami innych skupień. Następnie obliczono ich postacie sumaryczne, tj. średnie arytmetyczne z mierników cząstkowych, w oparciu o które wyznaczono ich ilorazy. Dodatkowo, w celu określenia odpowiedniej liczby klas posłużono się miarami opartymi na macierzach wariancji wewnątrzgrupowej i międzygrupowej, tj. obliczono ślady tych macierzy oraz ich iloraz [Grabiński i in. 1989, 155-157]. W oparciu o wartości śladów wspomnianych dwóch macierzy obliczony został także indeks Calińskiego i Harabasza ( $C-H$ ), [Calinski i Harabasz 1974, Parysek 1982, 108].

rezultaty potwierdzają wcześniejszy wniosek o demograficznym odmłodzeniu subpopulacji legitymującej się wyższym poziomem wykształcenia (zob. Tabela 4).

Obszary o najniższej medianie wieku ludności omawianej kategorii wykształcenia w obu latach zlokalizowane były w środkowej i wschodniej części kraju, o najwyższej zaś w dużych i wielkich miastach, a także na obszarach je okalających, a w roku 2011 również w powiatach położonych w północnej części Polski (zob. Rysunek 3 i 4).

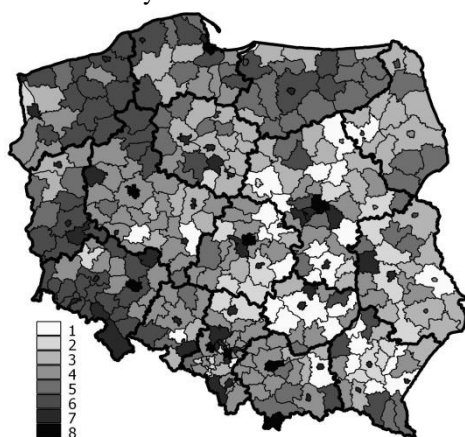
Tabela 4. Udziały poszczególnych grup wieku (w %) w subpopulacji posiadającej wykształcenie wyższe w skupieniach powiatów wyodrębnionych metodą *k*-średnich w latach 2002 i 2011 oraz mediana wieku ludności (w latach)

Grupy wieku	Numer skupienia – 2002 r.								Numer skupienia – 2011 r.							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
20-29	33,5	31,6	27,7	27,9	24,1	24,1	24,2	20,3	37,7	38,5	34,3	29,8	31,3	24,2	26,7	23,1
30-39	26,0	26,9	27,0	24,5	26,0	23,6	22,4	19,6	31,3	27,1	29,5	32,8	27,8	35,2	29,3	28,8
40-49	20,3	20,1	22,7	22,0	23,7	24,7	22,0	21,1	14,0	15,6	15,0	15,3	16,1	17,7	16,3	14,6
50-59	11,6	12,6	13,1	14,9	14,9	16,4	17,1	18,6	9,4	10,4	11,5	11,9	13,4	12,4	14,2	15,0
60+	8,7	8,9	9,5	10,8	11,3	11,2	14,3	20,4	7,5	8,5	9,8	10,3	11,4	10,6	13,4	18,4
Razem	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Me	35,9	36,1	37,6	38,2	39,0	40,7	41,3	45,1	32,5	32,7	33,4	35,2	35,7	36,6	37,0	38,3

Źródło: GUS [2015], opracowanie własne

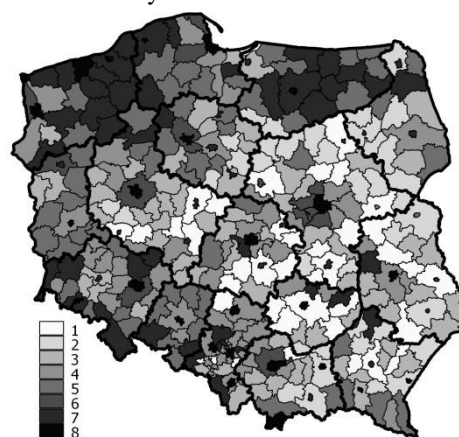
Rysunek 3. i Rysunek 4. Skupienia powiatów uzyskane metodą *k*-średnich z punktu widzenia struktury wieku ludności posiadającej wykształcenie wyższe w latach 2002 i 2011

Rysunek 3 – rok 2002



Źródło: GUS [2015], opracowanie własne

Rysunek 4 – rok 2011



Źródło: GUS [2015], opracowanie własne

W celu porównania podobieństwa przeprowadzonych klasyfikacji dokonano zestawienia w tablicy kontyngencji powiatów według ich przynależności do poszczególnych skupień. Z informacji tych wynika, że 138 jednostek (tj. 36,4% ogółu powiatów) w 2011 roku nie zmieniło swojej przynależności do skupienia do którego zostało przyporządkowane w roku 2002 (najwięcej w grupie nr 4), natomiast 115 powiatów (30,3%) zostało sklasyfikowane w grupie starszej, a 126 jednostek (33,2%) w grupie młodszej (biorąc pod uwagę medianę wieku ludności). Potwierdza to wcześniejszy wniosek o demograficznym odmłodzeniu się w rozpatrywanym okresie subpopulacji legitymującej się wyższym poziomem wykształcenia.

Następnie obliczony został wskaźnik zgodności<sup>11</sup>, którego wartość wynosząca 0,36 świadczy o relatywnie niskiej zgodności czasowo-przestrzennej obu klasyfikacji.

#### KLASYFIKACJE WOJEWÓDZTW I POWIATÓW Z PUNKTU WIDZENIA STRUKTURY PŁCI LUDNOŚCI LEGITYMUJĄCEJ SIĘ WYKSZTAŁCENIEM WYŻSZYM W LATACH 2002 i 2011

Następnie, stosując metodę *k*-średnich, przeprowadzone zostały próby klasyfikacji powiatów z punktu widzenia struktury płci ludności posiadające wyższy poziom wykształcenia w latach 2002 i 2011. W tym przypadku także dokonano kolejno podziałów na 4-10 skupień, a uzyskane rezultaty zostały poddane ocenie merytorycznej, przy wykorzystaniu ww. mierników jakości klasyfikacji.

Spośród przeprowadzonych podziałów powiatów za najbardziej odpowiedni w przypadku obu lat uznano podział na siedem skupień. Utworzone grupy następnie zostały uporządkowane niemalejąco według różnicy pomiędzy udziałami kobiet i mężczyzn (w pkt. proc.). Uzyskane rezultaty wskazują na wzrost

<sup>11</sup> Wskaźnik zgodności *W* [zob. Nowak 1990, 138] obliczany jest w oparciu o zestawione w macierzy *Z* o wymiarach *p* x *q* wyniki dwóch klasyfikacji (tj. przynależność obiektów do poszczególnych skupień) i ma postać:

$$w = \frac{1}{n_p + n_q} \left( \sum_{p=1}^p z_{p0} + \sum_{q=1}^q z_{0q} \right)$$

$$\text{gdzie: } z_{p0} = \max_q \{z_{pq}\}, \quad z_{0q} = \max_p \{z_{pq}\}, \quad z_{pq} = \frac{n_{pq}}{\max(n_p, n_q)}$$

$n_{pq}$  - liczba obiektów przyporządkowanych jednocześnie do grup  $G_p$  i  $G_q$ ;

$n_p, n_q$  - to liczby obiektów przyporządkowanych odpowiednio do grup  $G_p$  i  $G_q$ ,

$n_p + n_q + n_{pq} = N$

Wskaźnik ten przyjmuje wartości z przedziału  $[1/N; 1]$  – wyższa jego wartość świadczy o większym podobieństwie rezultatów porównywanych klasyfikacji.



dysproporcji pomiędzy udziałami kobiet i mężczyzn w subpopulacji legitymującej się wyższym poziomem wykształcenia (zob. Tabela 5).

Najniższe różnice punktowe pomiędzy płciami w obrębie omawianej kategorii wykształcenia w obu latach (skupienia nr 1 i 2) zaobserwowane zostały w dużych miastach oraz na ich obrzeżach (przy czym, obszary te legitymowały się relatywnie wysokim udziałem osób z wykształceniem wyższym) oraz w zachodniej i południowo-zachodniej części kraju (przy czym, obszary te cechowały się względnie niskim odsetkiem omawianej subpopulacji). Natomiast najwyższymi różnicami punktowymi (skupienia nr 6 i 7) charakteryzowały się powiaty zlokalizowane na obrzeżach województw, w szczególności w środkowej i wschodniej części Polski.

Tabela 5. Udziały kobiet i mężczyzn (w %) w subpopulacji posiadającej wykształcenie wyższe w skupieniach powiatów wyodrębnionych metodą *k*-średnich w latach 2002 i 2011 oraz różnice (w pkt. proc.) pomiędzy tymi udziałami (K-M)

	Numer skupienia – 2002 rok							Numer skupienia – 2011 rok						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Mężczyźni	48,5	45,9	44,0	42,1	40,3	38,4	36,3	44,6	42,3	40,9	39,8	38,7	37,6	35,9
Kobiety	51,5	54,1	56,0	57,9	59,7	61,6	63,7	55,4	57,7	59,1	60,2	61,3	62,4	64,1
Razem	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Źródło: GUS [2015], opracowanie własne

Porównując uzyskane klasyfikacje można stwierdzić, że 145 jednostek (tj. 38,3% ogółu powiatów) w 2011 roku nie zmieniło swojej przynależności do skupienia do którego zostało przyporządkowane w roku 2002 (najwięcej w grupie nr 4), natomiast 153 powiatów (40,4%) zostało sklasyfikowane w grupie o wyższym numerze, a 81 jednostek (21,4%) w grupie o niższym numerze (biorąc pod uwagę różnice punktowe w odsetkach frakcji kobiet i mężczyzn). Klasyfikacje te charakteryzują się relatywnie niską zgodnością rezultatów (wskaźnik zgodności wyniósł 0,38).

## PODSUMOWANIE

W Polsce w latach 2002-2011 przemianom uległa struktura wykształcenia ludności. Zmiany te objawiały się przede wszystkim wzrostem udziału ludności legitymującej się wykształceniem wyższym i spadkiem frakcji posiadających niższe poziomy wykształcenia.

Analiza subpopulacji legitymującej się wyższym poziomem wykształcenia z punktu widzenia struktur według płci i wieku ludności wskazała na brak stabilności w czasie tych charakterystyk. W rozpatrywanym okresie odnotowano spadek mediany wieku ludności posiadającej omawianą kategorię wykształcenia, zwiększeniu natomiast uległa dysproporcja pomiędzy udziałami kobiet i mężczyzn w rozpatrywanej subpopulacji. Opisywane przemiany widoczne były także

w przekrojach terytorialnych, o czym świadczą rezultaty przeprowadzonych klasyfikacji województw i powiatów.

## BIBLIOGRAFIA

- Calinski T., Harabasz J. (1974) A dendrite method for cluster analysis, *Communications in Statistics*, Vol. 3, No. 1, pp. 1–27.
- Grabiński T., Grabiński T., Wydymus S., Zeliaś A. (1989) *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*, PWN, Warszawa.
- GUS (2015) Bank Danych Lokalnych, Główny Urząd Statystyczny, [http://stat.gov.pl/bdl/app/strona.html?p\\_name=indeks](http://stat.gov.pl/bdl/app/strona.html?p_name=indeks) (dostęp 29.06.2015).
- GUS (2013) *Ludność. Stan i struktura demograficzno-społeczna. Narodowy Spis Powszechny Ludności i Mieszkań 2011*, Główny Urząd Statystyczny, Warszawa.
- Młodak A. (2006) *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa.
- MNiSW (2015) Proces boloński, <http://www.nauka.gov.pl/proces-bolonski/proces-bolonski.html> (dostęp 29.06.2015).
- Nowak E. (1990) *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWN, Warszawa.
- Panek T. (2009) *Statystyczne metody wielowymiarowej analizy porównawczej*, SGH, Warszawa.
- Parysek J., J. (1982) *Modele klasyfikacji w geografii*, Seria Geografia, nr 31, Uniwersytet im. Adama Mickiewicza w Poznaniu, Poznań.
- Ustawa z dn. 27 lipca 2005 r., Prawo o szkolnictwie wyższym, Dz.U. 2005 nr 164 poz. 1365.

## CLASSIFICATION OF POLISH AREAS FROM THE PERSPECTIVE OF CHOSEN SOCIO-DEMOGRAPHIC CHARACTERISTICS

**Abstract:** The paper aims at assessing the similarity between Polish regions in terms of selected socio-demographic characteristics. Focus will be placed on the education structure of the population, within which I will consider the age and sex structure of the population with higher education. The survey will be conducted at the level of voivodeships and poviats in 2002 and 2011, using selected methods of multidimensional comparative analysis (including Ward method and *k*-means method). I will also investigate the dynamics of changes in these structures over a given period. The analysed data will be obtained from the publications of the Central Statistical Office on the structure of education of the population, derived from the results of population censuses, carried out in Poland in 2002 and 2011.

**Keywords:** the education structure of the population, the age and sex structure of the population with higher education, multidimensional comparative analysis