

*RAFAŁ KLÓSKA, RAFAŁ CZYŻYCKI*

**WIELOWYMIAROWA ANALIZA STATYSTYCZNA  
POZIOMU ROZWOJU DEMOGRAFICZNEGO SZCZECINA  
NA TLE INNYCH MIAST WOJEWÓDZKICH W POLSCE**

**Uwagi metodyczne**

Celem artykułu jest statystyczna analiza poziomu rozwoju demograficznego miast wojewódzkich w Polsce<sup>1</sup>. Kolejność poszczególnych miast wojewódzkich ustalono z uwagi na przyjęte kryterium ogólne oraz przeprowadzono analizę skupień w celu wyodrębnienia podgrup podobnych aglomeracji.

Zauważmy, że rozwój demograficzny - stanowiący przedmiot analizy - jest pewną wielowymiarową charakterystyką, której sposób mierzenia pozostaje niejednoznaczny. Ze względu na nieprecyzyjność tego typu terminów najczęściej zakłada się daleko idącą powszechną zgodę co do ich znaczenia. Wyniki badań determinuje w głównej mierze ostateczna lista zmiennych diagnostycznych, jak również wybór miary odległości i metody grupowania (przy grupowaniu) oraz formuły agregacji przy porządkowaniu liniowym). W związku z istnieniem wielu formuł normalizacji zmiennych, sposobów określania wag, metod uśredniania wartości znormalizo-

---

<sup>1</sup> W Raporcie *Gospodarka Szczecina 2007. Diagnoza i perspektywy rozwoju w ujęciu krótko-średniookresowym*, wykonanym przez zespół pod kierunkiem prof. zw. dr. hab. S. Flejterskiego na zlecenie Gminy Miasto Szczecin, reprezentowanej przez Wydział Rozwoju Miasta i Funduszy Pomocowych, przedstawiono rezultaty badań poziomu rozwoju społeczno-gospodarczego Szczecina na tle innych miast wojewódzkich w Polsce. Podjęte wówczas badania są kontynuowane, a niniejszy artykuł zawiera wyniki pozwalające ocenić stopień zróżnicowania demograficznego aglomeracji w Polsce.

wanych, sposobów ustalania współrzędnych obiektu odniesienia i formuł obliczania odległości w literaturze naukowej opisano wiele różnych miar agregatowych (wykorzystywanych w praktyce przy m.in. sporządzaniu różnego typu rankingów). Należy zatem pamiętać, że różne formuły agregacji mogą dawać różne wyniki końcowe, nawet w odniesieniu do kryterium ogólnego, reprezentowanego przez tę samą listę zmiennych diagnostycznych. Ostateczna lista zmiennych uwzględnionych w badaniu ma jednak decydujące znaczenie przy klasyfikacji obiektów, a więc powinna być przedyskutowana w gronie ekspertów i uznana za najlepiej reprezentującą analizowane zagadnienie.

Z uwagi na przesłanki merytoryczne oraz dostępność danych statystycznych jako zmienne determinujące rozwój demograficzny uwzględniono - przedyskutowane w gronie ekspertów - następujące cechy statystyczne<sup>2</sup>:

$X_1$  - urodzenia żywe na 1000 mieszkańców,

$X_2$  - zgony na 1000 mieszkańców,

$X_3$  - przyrost naturalny na 1000 mieszkańców,

$X_4$  - małżeństwa na 1000 mieszkańców,

$X_5$  - saldo migracji wewnętrznych i zagranicznych na pobyt stały na 1000 mieszkańców,

$X_6$  - zgony niemowląt na 1000 urodzeń żywych,

$X_7$  - ludność na 1 km<sup>2</sup>.

Ilustrując graficznie analizowane dane wielowymiarowe, wykorzystano wielowymiarowe wykresy obrazkowe jako jedne z lepszych ogólnych technik eksploracyjnej analizy danych. *Twarze Chernoffa* pozwalają przedstawić obserwacje wielowymiarowe w postaci zarysów ludzkich twarzy, dzięki czemu podobieństwo analizowanych w badaniu miast wojewódzkich można oceniać na podstawie podobieństwa twarzy zdefiniowanych za pomocą branych pod uwagę siedmiu cech statystycznych.

Ustalając ranking miast wojewódzkich w Polsce ze względu na poziom rozwoju demograficznego opisywany przez cechy uwzględnione w badaniu, wykorzystano, wchodzący w skład metod porządkowania liniowego, *względny współczynnik rozwoju*<sup>3</sup>, czyli miarę agregatową stanowiącą średnią arytmetyczną ze zmien-

---

<sup>2</sup> Dane statystyczne obrazują stan na koniec roku 2006, a pochodzą z *Rocznika Statystycznego Województw 2007*, Główny Urząd Statystyczny, Warszawa 2007, s. 190-211.

<sup>3</sup> Por. A. Sokołowski, *Analizy wielowymiarowe. Materiały kursowe StatSoft Polska*, Kraków 2005, s. 19-20.

nych diagnostycznych sprowadzonych do porównywalności poprzez unitaryzację<sup>4</sup>, wyrażoną wzorem:

$$W_i = \frac{100}{m} \sum_{j=1}^m \alpha_j x'_{ij},$$

gdzie:

- $W_i$  - względny współczynnik rozwoju,
- $m$  - liczba cech statystycznych branych pod uwagę w badaniu,
- $\alpha_j$  - waga  $j$ -tej zmiennej,
- $x_{ij}$  - znormalizowane metodą unitaryzacji wartości cech statystycznych branych pod uwagę w badaniu, przy czym algorytm unitaryzacji jest następujący:

dla stymulant:

i

dla destymulant:

$$x'_{ij} = \frac{x_{ij} - \min\{x_{ij}\}}{\max\{x_{ij}\} - \min\{x_{ij}\}};$$

$$x'_{ij} = \frac{\max\{x_{ij}\} - x_{ij}}{\max\{x_{ij}\} - \min\{x_{ij}\}}.$$

Względny współczynnik rozwoju  $W_i$  przyjmuje wartości z przedziału od 0 do 100, przy czym im wyższa wartość tego miernika, tym wyższe miejsce analizowanego obiektu (tu: miasta) w rankingu. Identyfikując charakter każdej z siedmiu zaproponowanych w badaniu zmiennych ze względu na ich wpływ na rozwój demograficzny, przyjęto, że zmienne opisujące zgony ( $X_2$  i  $X_6$ ) to destymulanty, a pozostałe cechy to stymulanty.

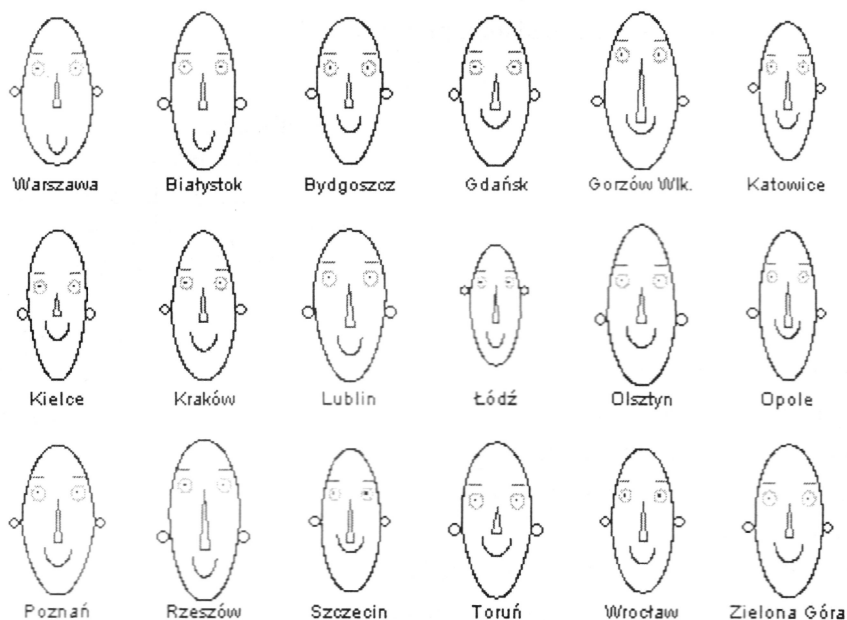
Poszukując skupisk miast wojewódzkich o podobnym poziomie rozwoju demograficznego zastosowano aglomeracyjną metodę Warda z wykorzystaniem odległości euklidesowej odnośnie branych pod uwagę zmiennych diagnostycznych, sprowadzonych do porównywalności metodą standaryzacji. Podejmując próbę weryfikacji uzyskanych wyników zastosowano jednoczynnikową analizę wariancji. W celu sprawdzenia hipotezy zerowej o całkowitym braku oddziaływania czynnika klasyfikacyjnego (wpływu każdej ze zmiennych diagnostycznych) na rezultaty grupowania wykorzystano test istotności F oparty na rozkładzie Fishera-Snedecora.

<sup>4</sup> Unitaryzacja, standaryzacja i rangowanie to trzy najczęściej dziś stosowane metody normalizacji zmiennych.

Decyzją optymalną jest zatem odrzucenie hipotezy zerowej na zadeklarowanym poziomie istotności  $\alpha$ , przy czym w badaniu z góry przyjęto  $\alpha = 0,05$ .

### Porządkowanie liniowe miast wojewódzkich w Polsce ze względu na poziom rozwoju demograficznego

Ilustrując graficznie analizowane dane statystyczne<sup>5</sup> z roku 2006 wykorzystano wykres obrazkowy *Twarze Chernoffa*<sup>6</sup> (rys. 1). Można zauważyć, że Szczecin ze względu na analizowane cechy różni się od pozostałych aglomeracji.



Rys. 1. Twarze Chernoffa

Źródło: opracowanie własne.

<sup>5</sup> Ze względu na fakt opisywania każdego z 18 analizowanych miast wojewódzkich przy wykorzystaniu siedmiu tych samych cech mamy do czynienia z 18 obiektami w przestrzeni siedmiowymiarowej.

<sup>6</sup> Przeporządkowanie cech poszczególnym elementom twarzy przyjęto domyślnie - proponowane przez program STATISTICA firmy StatSoft, przy użyciu którego wykonano wszystkie niezbędne obliczenia na potrzeby niniejszego opracowania (por. A. Luszniwicz, T. Słaby, *Statystyka z pakietem komputerowym STATISTICA PL. Teoria i zastosowania*, Wydawnictwo C.H. Beck. Warszawa 2003).

Analizując prezentowane w tabeli 1 wartości względnego współczynnika poziomu rozwoju demograficznego  $W_i$  otrzymane w wyniku nieważonych średnich wartości znormalizowanych zmiennych  $X_1, X_2, \dots, X_7$  pomnożonych przez 100, można zauważyć, że w 2006 roku Szczecin na tle innych miast wojewódzkich w Polsce plasuje się na 15. miejscu<sup>7</sup>.

Tabela 1

Ranking miast wojewódzkich na podstawie wartości względnego wskaźnika poziomu rozwoju demograficznego w roku 2006

Miasto	Miejsce w rankingu	Wartość $W_i$
Warszawa	1	71
Białystok	2	68
Olsztyn	3	<b>66</b>
Rzeszów	4	<b>66</b>
Lublin	5	64
Zielona Góra	6	64
Toruń	7	62
Kraków	8	57
Gdańsk	9	57
Poznań	10	54
Bydgoszcz	11	54
Kielce	12	53
Gorzów Wlkp.	13	51
Wrocław	14	50
<b>Szczecin</b>	15	40
Opole	16	37
Katowice	17	35
Łódź	18	20

Źródło: obliczenia i opracowanie własne.

Należy jednak podkreślić, że wartości zastosowanego miernika  $W_i$  są w umiarkowanym stopniu zróżnicowane, a typowe miasto wojewódzkie miało względny współczynnik poziomu rozwoju demograficznego  $W_i$  w przybliżonych granicach od 40 do 67. Przyjmując skalę od 0 do 100% różnica wartości zastoso-

<sup>7</sup> Analizując różnego typu rankingi należy mieć świadomość, że różne formuły agregacji stosowane do oceny merytorycznego przedmiotu analizy opisywanego przez zmienne diagnostyczne, których wybór często pozostaje niejednoznaczny, mogą dawać różne wyniki końcowe.

wanego miernika poziomu rozwoju między Szczecinem a Warszawą - zwycięzcą rankingu - wyniosła 31 punktów procentowych, przy różnicy między Szczecinem a Łodzią - przegranym rankingu - na poziomie 20 punktów procentowych.

### **Grupowanie miast wojewódzkich w Polsce ze względu na poziom rozwoju demograficznego**

Poszukując skupisk miast wojewódzkich o podobnym poziomie rozwoju demograficznego zastosowano aglomeracyjną metodę Warda z wykorzystaniem odległości euklidesowej. Na podstawie przyjętych do badania siedmiu zmiennych diagnostycznych, sprowadzonych do porównywalności metodą standaryzacji, otrzymano drzewko połączeń dla roku 2006. Analizując odpowiedni dendrogram, rozsądny wydawał się podział miast wojewódzkich Polski na pięć grup, przy czym podejmując próbę weryfikacji uzyskanych wyników zastosowano jednoczynnikową analizę wariancji. Na przyjętym poziomie istotności 0,05 jedynie liczba ludności na 1 km<sup>2</sup> ( $X_7$ ) okazała się nieistotna statystycznie. Tym samym pozostałe zmienne diagnostyczne uwzględnione w badaniu w sposób istotny statystycznie wpłynęły na pogrupowanie miast wojewódzkich w skupienia o podobnym poziomie rozwoju demograficznego (rys. 2 i tab. 2):

Grupa I - Łódź

Grupa II - Wrocław, Szczecin, Opole, Katowice

Grupa III - Rzeszów, Gorzów Wlkp.

Grupa IV - Toruń, Kielce, Poznań, Zielona Góra, Gdańsk, Lublin, Bydgoszcz

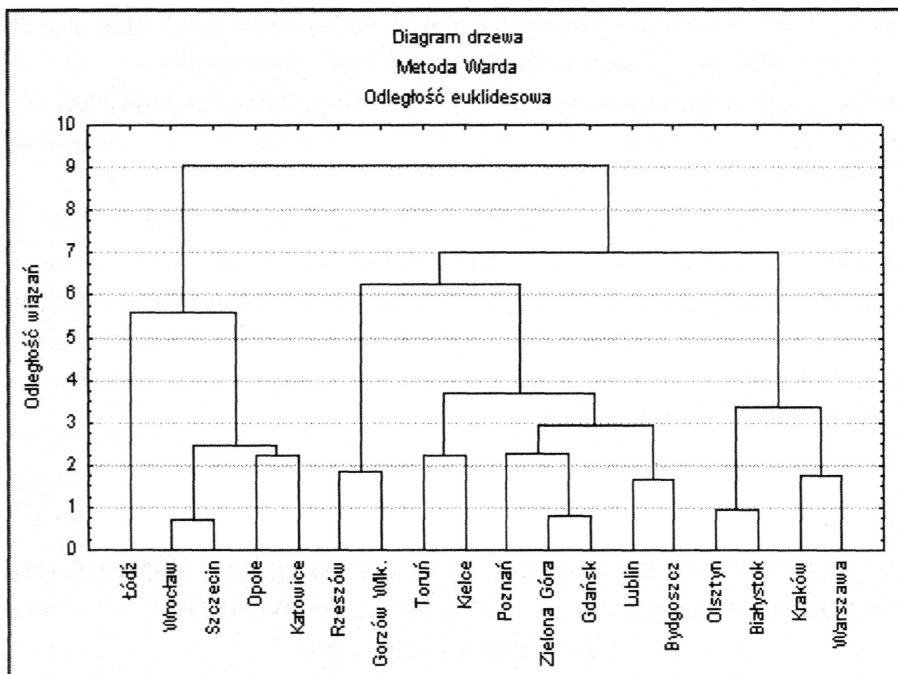
Grupa V - Olsztyn, Białystok, Kraków, Warszawa

Tabela 2

Wybrane wyniki jednoczynnikowej analizy wariancji

Zmiana	Analiza wariancji			
	Zaznaczone efekty są istotne z $p < 0,05$			
$x_1$	0,007464			
$x_2$	0,001553			
$x_3$	0,000054			
$x_4$	0,000392			
$x_5$	0,035356			
$x_6$	0,001097			

Źródło: obliczenia i opracowanie własne.



Rys. 2. Wyniki grupowania

Źródło: obliczenia i opracowanie własne.

Na podstawie otrzymanych wyników można zatem stwierdzić, że na koniec roku 2006 Szczecin pod względem poziomu rozwoju demograficznego, opisywane-go przez cechy ujęte w badaniu, był najbardziej podobny do Wrocławia.

### Bibliografia (wybór)

- Kłóska R., *Wielowymiarowa analiza statystyczna rozwoju społeczno-gospodarczego Szczecina na tle innych miast wojewódzkich w Polsce*. W: *Region 2008. Pozyskiwanie inwestorów a rozwój regionalny*, Uniwersytet Szczeciński, Szczecin 2008.
- Kolenda M., *Taksonomia numeryczna. Klasyfikacja, porządkowanie i analiza obiektów wielocechowych*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2006.
- Luszniewicz A., Słaby T., *Statystyka z pakietem komputerowym STATISTICA PL. Teoria i zastosowania*, Wydawnictwo C.H. Beck, Warszawa 2003.

- Malina A., *Analiza czynnikowa jako metoda klasyfikacji regionów Polski*, „Przegląd Statystyczny” 2006, nr 1, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Młodak A., *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa 2006.
- Pociecha J., Podolec B., Sokołowski A., Zając K., *Metody taksonomiczne w badaniach społeczno-ekonomicznych*, PWN, Warszawa 1988.
- Rocznik Statystyczny Województw 2007*, Główny Urząd Statystyczny, Warszawa 2007.
- Sokołowski A., *Analizy wielowymiarowe*, Materiały kursowe Statsoft Polska, Kraków 2005.
- Walesiak M.: *Statystyczna analiza wielowymiarowa w badaniach marketingowych*, Prace Naukowe Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1993.
- Zeliaś A. (red.), *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*, PWN, Warszawa 1989.

## THE MULTIVARIATE STATISTICAL LEVEL ANALYSIS OF DEMOGRAPHIC DEVELOPMENT OF SZCZECIN IN COMPARISON WITH OTHER VOIVODSHIP CITIES IN POLAND

### Summary

Quantitative methods find a great application in analysis processes, diagnoses and economic prognoses by means of description and estimation of forming economic variables in time and space as well as expectations regarding direction and character of changes of these variables are becoming more precise. In the article the level of demographic development of Polish voivodship cities has been analysed. The sequence of individual provinces has been established considering a stated general criterion represented by the suggested variables in the research and cluster analysis has been carried out to isolate subgroups of similar provinces.