

Analiza zjawisk społeczno-ekonomicznych z zastosowaniem metod taksonomicznych

Wyodrębnienie wspólnej charakterystyki może posłużyć do określenia różnic między badanymi grupami jednostek i jednocześnie określić ich wspólne cechy w obrębie tejże grupy. Istnieje wiele metod statystycznych pozwalających dokonać takiego podziału, a na szczególne miejsce, ze względu na prostotę i popularność, zasługują metody taksonomiczne. W wielu prowadzonych analizach z ich wykorzystaniem nie przestrzega się jednak podstawowych zasad, jakie muszą spełniać zmienne wykorzystywane w prowadzonych opisach zjawisk oraz kryteriów stosowania określonej metody. Od poprawności ich użycia zależy jakość uzyskanych wyników.

Artykuł poświęcony jest analizie porównawczej wniosków uzyskanych przy opisie przestrzennego zróżnicowania zjawisk społeczno-ekonomicznych z wykorzystaniem różnych metod taksonomicznych oraz wpływu różnych ich własności na wyniki. Przedstawiony będzie też ich syntetyczny opis, ze szczególnym uwzględnieniem warunków, jakie muszą spełniać zmienne w prowadzonych opisach zjawisk.

ZAKRES ZASTOSOWANIA METOD TAKSONOMICZNYCH

Podstawowym wymogiem stosowania metod taksonomicznych jest określenie obiektów, które będą ze sobą porównywane oraz cech, które będą określać ich właściwości. W analizach przestrzennego zróżnicowania¹ są to zazwyczaj wyodrębnione w sposób naturalny lub sztuczny jednostki², tzn. kontynenty, kraje,

¹ Badania nad przestrzennym zróżnicowaniem procesów społeczno-gospodarczych rozpoczęto na początku XIX w. i dotyczą prac D. A. Ricarda i J. H. Thunena. Zauważyli oni, że istnieje zależność między efektem produkcji a odległością od rynku zbytu. Problemem lokalizacji przemysłu w sensie mikroekonomicznym (określenie optymalnego usytuowania w przestrzeni pojedynczego przedsiębiorstwa lub gospodarstwa) oraz makroekonomicznym (przestrzenne rozmieszczenie całej badanej działalności gospodarczej) zajmował się również A. Weber. W badaniach wykorzystał po raz pierwszy w szerokim zakresie metody ilościowe. Nad zagadnieniami teorii lokalizacji przemysłu i rolnictwa w ujęciu dynamicznym, opartymi na formułach matematycznych, pracował A. Losch. Powstało również wiele badań o charakterze przestrzennym, związanych z problemem natury ekologicznej, społeczno-ekonomicznej, technicznej czy antropologicznej. Należy tu wymienić chociażby prace J. J. Rousseau, B. D. Sain Pierre (Nowak, 1990).

² Początki wyodrębnienia jednorodnych struktur przestrzennych możemy odnaleźć w próbach określenia fizycznych kryteriów różnicowania ziemi, które pozwoliły na wyodrębnienie naturalnych krain geograficznych. Za pierwsze dzieło zajmujące się tą problematyką uznaje się *Wstęp do geografii* Ptolemeusza z I wieku n.e., zaś twórcą geografii stosowanej jest grecki geograf Strabon — autor podręcznika „Geografia” (*Encyklopedia...*, 1982).

regiony, podregiony, województwa, powiaty czy gminy. Badanym obiektem może być również osoba, populacja, przedsiębiorstwo czy szkoła. Są one opisywane za pomocą różnych zmiennych, których wybór zależy od celu prowadzonych analiz. Na przykład mogą one opisywać warunki życia ludności, poziom rozwoju rolnictwa, efektywność przedsiębiorstw, stan zdrowia ludności, infrastrukturę w ochronie zdrowia, sytuację na rynku pracy czy mieszkalnictwo. Analizy te prowadzone są w ujęciu statycznym lub dynamicznym, a wyodrębniania grup obiektów podobnych można było dokonywać poprzez podział badanej zbiorowości na jednorodne podgrupy (analiza z góry) lub dołączać do kolejnych obiektów obiekty podobne (analiza z dołu). O popularności stosowania tych metod świadczy częstotliwość ich wykorzystywania.

Szerokie wykorzystanie metod taksonomicznych w badaniach prowadzonych w różnych dziedzinach nauki każe się zastanowić, czy zawsze są one poprawnie stosowane. Czy analizowane zmienne spełniają stawiane im wymogi formalne i statystyczne oraz czy dobór metody do opisu zjawiska jest właściwy. Od spełnienia tych warunków zależy bowiem poprawność uzyskanych wyników.

ZARYS TEORII METOD TAKSONOMICZNYCH

Analizy porównawcze obejmujące statystyczne, ekonometryczne i ekonomiczne rozważania o rynku, działalności przedsiębiorstw, procesach demograficzno-społecznych oparte są zazwyczaj na modelach uwzględniających małą liczbę zmiennych. Jednocześnie większość zjawisk opisywana jest przez wiele różnych zmiennych (ilościowych i jakościowych). Konieczna zatem staje się ich agregacja (Pluta, 1977).

Decyzje podejmowane w państwie w skali makro czy przez szefów dużych przedsiębiorstw odnoszą się nie do każdej z jednostek oddzielnie, lecz do całych ich zbiorowości, przy jednoczesnym uwzględnieniu specyfiki tych obiektów. Prowadzenie syntetycznych analiz poziomu i struktury omawianych obiektów jest zatem niezbędne (Grabiński i in., 1976), ale równocześnie niezwykle trudne. Z jednej strony dotyczy bowiem złożonych zjawisk społeczno-gospodarczych, z drugiej zaś — dużej liczby obiektów i ich różnej specyfiki.

Wyodrębnienie homogenicznych grup obiektów umożliwia przeprowadzenie pogłębionych analiz w obrębie tychże grup przy uwzględnieniu większej liczby zmiennych. Podejście to pozwala na lepsze poznanie czynników decydujących o poziomie i strukturze badanego zjawiska, a w konsekwencji na trafniejszą ocenę istniejącej rzeczywistości oraz wyodrębnienie ewentualnych przyczyn różnicujących porównywane obiekty (Gorzelać, 1981). Najczęściej analizy takie dotyczą porównań między obiektami w określonym momencie, tych samych obiektów w różnych punktach czasowych lub porównań z obiektem realnie nieistniejącym — obiektem wzorcem (Nowak, 1990).

Istnieje wiele różnych podziałów metod taksonomicznych. Pierwsza z nich za główne kryterium uznaje **cel prowadzonych rozważań** (Styczeń, 1971; Kowal-

ski, 1977; Bartosiewicz, 1976; Pluta, 1977). Wyróżnia się tu *taksonomię porządkującą*, która obejmuje metody porządkujące badane obiekty. *Metoda liniowa* polega na rzutowaniu punktów przestrzeni wielowymiarowej na prostą (np. metoda Czekanowskiego). *W metodzie nieliniowej* punkty rzutowane są na płaszczyznę (np. metoda dendrytów). *Taksonomia podziałowa* dotyczy podziału zbioru elementów na jednorodne grupy (np. metoda kul). *Taksonomia wyboru reprezentanta grup* stosowana jest w celu wybrania reprezentantów grup.

Kolejny podział opiera się na **sposobie prezentacji uzyskanych wyników**. Istotą *taksonomii graficznej* jest forma prezentacji (np. diagramy, dendryty, dendrogramy), zaś *analitycznej* — brak ilustracji graficznej.

Są też **metody uwzględniające przekształcenia zmiennych** ujętych w klasyfikacji. *Taksonomia czynnikowa* stosuje przekształcenia ortogonalne macierzy danych wyjściowych, a *taksonomia nieczynnikowa* — przekształcenia nieortogonalne.

Inną klasyfikację metod taksonomicznych przeprowadził T. Grabiński (Grabiński i in., 1991). Dokonał on ich podziału ze względu na następujące warunki:

- ♦ *rozpoczęcie procesu grupowania* (metody aglomeracyjne — początkowo każdy obiekt jest odrębną grupą, która łączy się następnie z inną grupą, proces ten trwa aż do uzyskania jednej grupy; metody podziałowe — początkowo cały zbiór obiektów jest jedną grupą, która dzieli się na podgrupy aż do uzyskania grup jednoelementowych);
- ♦ *hierarchia grup* (metody hierarchiczne — każda grupa wyższego rzędu zawiera w sobie grupy niższego rzędu; niehierarchiczne — grupy niższego rzędu nie muszą wchodzić w skład grup wyższego rzędu);
- ♦ *rozłączność grup* (metody rozłączne — grupy nie zachodzą na siebie; metody nierozłączne — grupy mogą zachodzić na siebie);
- ♦ *sposób grupowania* (metody sekwencyjne — istnieją powtarzające się ciągi operacji; metody równoczesne — brak powtarzających się ciągów operacji);
- ♦ *kryterium grupowania* (metody globalne — na każdym etapie grupowania wykorzystuje się to samo kryterium optymalizacji; metody lokalne — optymalizacja dokonywana jest odrębnie na każdym etapie grupowania);
- ♦ *algorytm grupowania* (metody bezpośrednie — grupowanie przebiega w jednym ciągu operacji; metody iteracyjne — algorytm tworzą kolejne iteracje);
- ♦ *wagi grup* (metody ważone — znaczenie grup obiektów lub wymiarów przestrzeni klasyfikacji jest zróżnicowane; metody nieważone — jednakowa waga grup obiektów lub wymiarów przestrzeni klasyfikacji);
- ♦ *możliwość uczenia się* (metody adaptacyjne — algorytm ma możliwość zmiany zasad grupowania w zależności od uzyskiwanych wyników; metody nieadaptacyjne — zasady grupowania w procedurze są stałe).

Z przedstawionych przykładów sposobów klasyfikacji metod taksonomicznych jasno wynika, że nie są one rozłączne. Te same metody mogą należeć jednocześnie do wielu wyodrębnionych grup, zależnie od przyjętego kryterium podziału.

W rozważaniach prowadzonych z wykorzystaniem metod taksonomicznych, podobnie jak i w innych procedurach statystycznych, przed ich rozpoczęciem muszą być podjęte decyzje. Dotyczą one takich zagadnień, jak: wybór zmiennych wykorzystanych w analizie, miara odległości, kryteria, jakie zostaną przyjęte przy formowaniu grup o podobnych własnościach. Czynniki te wpływają na wyniki badań. **Brak logicznego powiązania między zmiennymi a opisywanym zjawiskiem spowoduje uzyskanie błędnych wyników. Przyjęcie różnych skal wartości zmiennych oznacza brak ich porównywalności. Prowadzić to może do nieproporcjonalnego odbicia ich udziału w wyznaczonej macierzy odległości.**

Podstawowymi pojęciami używanymi w teorii metod taksonomicznych są obiekt oraz cecha. Przez obiekt należy rozumieć jednostki analizy podlegające klasyfikacji. Wszystkie jednostki objęte rozważaniem tworzą zbiór obiektów, który możemy zapisać w następujący sposób: $\Omega = \{O_1, O_2, \dots, O_n\}$, gdzie n jest liczbą jednostek podlegających badaniu. Przez cechę (zmienną) należy rozumieć własność obiektów badanego zbioru, przy uwzględnieniu kryteriów klasyfikacji określonych przez cel badania. Zbiór zmiennych wykorzystywanych w opisie obiektów możemy zapisać jako: $X = \{X_1, X_2, \dots, X_k\}$. Zbiór informacji o k zmiennych dla n obiektów tworzy macierz obserwacji:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ x_{n1} & \dots & \dots & x_{nk} \end{bmatrix}$$

Należy pamiętać, że do zbioru zmiennych mogą należeć jedynie te, które w sposób logiczny powiązane są z opisywanym zjawiskiem. W sytuacji występowania cech niejednorodnych, zróżnicowanych pod względem poziomu wartości oraz jednostek pomiaru, istnieje konieczność sprowadzenia ich do porównywalności. Możemy tego dokonać za pomocą np. (Strahl, 1980):

- standaryzacji zmiennych $z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}$, przy czym: $i = \overline{1, n}$, $j = \overline{1, k}$

gdzie:

$\bar{x}_j$ — średnia arytmetyczna,

S_j — odchylenie standardowe wyznaczone dla określonej zmiennej

(zmienne wywierają jednakowy wpływ na odległości między badanymi jednostkami — jednakowa wariancja);

- normalizacji $z_{ij} = \frac{x_{ij}}{\max_i x_{ij}^2}$

(żadna ze zmiennych nie wywiera jednakowego wpływu na odległości między badanymi jednostkami — zróżnicowana wariancja);

- unitaryzacji $z_{ij} = \frac{x_{ij} - \min_i \{x_{ij}\}}{\max_i \{x_{ij}\} - \min_i \{x_{ij}\}}$

(przekształcone zmienne mają stały zakres zmienności).

W niektórych metodach taksonomicznych rozróżnia się wśród zmiennych zaproponowanych do analizy:

- stymulanty (wzrost wartości zmiennej wpływa korzystnie na analizowane zjawisko);
- destymulanty (wzrost wartości zmiennej wpływa niekorzystnie na analizowane zjawisko);
- nominanty (w pewnym przedziale zachowuje się jak stymulanta, a następnie jak destymulanta) (Pluta, 1986).

W badaniach taksonomicznych zasadnicze znaczenie ma dobór cech opisujących badane obiekty. Sprowadza się on do takiej redukcji zmiennych ze wstępnej listy potencjalnych zmiennych, aby pozostały tylko te, które odznaczają się jak **największą diagnostycznością badanego zjawiska**, tzn. zawierają istotne informacje o rozważanych procesach.

Dobór cech diagnostycznych musi spełniać określone kryteria merytoryczne i formalne, a także zapewniać odpowiednią wartość informacyjną zmiennych. Spełnienie kryterium merytorycznego oznacza, że cechy diagnostyczne muszą ujmować najbardziej istotne, a nie marginalne własności analizowanych obiektów, muszą być jednoznacznie i ściśle zdefiniowane oraz logicznie ze sobą powiązane. Powinny umożliwiać ich wzajemną kontrolę poprzez znajomość związków statystycznych i merytorycznych między nimi oraz charakteryzować się zgodnością proporcji między liczbą zmiennych reprezentujących dany aspekt zjawiska a ich znaczeniem merytorycznym.

Kryterium formalne wymaga, aby cechy diagnostyczne **były mierzalne** w sensie możliwości liczbowego wyrażenia ich poziomu. Informacje uzyskane powinny być wiarygodne i kompletne. Cechy winny charakteryzować się ciągłością, co czasami sprowadza się do konieczności eliminowania tych zmiennych, które uniemożliwiają wzajemną porównywalność obiektów zarówno w czasie, jak i w przestrzeni.

Kryterium wartości informacyjnej oznacza, że cechy diagnostyczne powinny mieć wysoką zdolność dyskryminacji obiektów, tzn. dużą zmienność w cza-

sie oraz charakteryzować się brakiem (lub niskim poziomem) korelacji między zmiennymi. Ma to na celu wyeliminowanie powtarzania informacji przez inne cechy. Silna korelacja powinna wystąpić między cechami diagnostycznymi i niewybranymi jako diagnostyczne (Grabiński i in., 1991).

Algorytm postępowania w opisie zjawisk z wykorzystaniem metod taksonomicznych polega na tym, że w pierwszym kroku procedury wybiera się te zmienne, które według posiadanej wiedzy merytorycznej stanowią najważniejsze cechy opisywanego zjawiska (zbiór zmiennych potencjalnych). W drugim kroku do oceny zdolności dyskryminacyjnych cech wykorzystuje się procedury statystyczno-matematyczne. Porównuje się wartości współczynników zmienności poszczególnych cech z arbitralnie przyjętą wartością krytyczną, a następnie ze zbioru potencjalnych cech diagnostycznych eliminuje się te, dla których wyznaczone wartości są mniejsze od wartości przyjętej. W celu oceny stopnia skorelowania cech najczęściej korzysta się ze współczynnika korelacji liniowej Pearsona.

Przedstawione kryteria wyboru cech diagnostycznych oparto na założeniu, że informacje o badanych obiektach i opisujących je cechach są kompletne. Niestety, nie zawsze tak jest. Brak danych powoduje powstanie luki w materiale statystycznym. Pojawia się problem badania w warunkach niepełnej informacji statystycznej. Jeżeli w materiale istnieją braki, należy przyjąć pewien dopuszczalny poziom ich tolerancji.

Z. Hellwig (1968) proponuje trzy poziomy tolerancji braku danych statystycznych: surowy poziom tolerancji — $p=10\%$, przeciętny poziom tolerancji — $p=20\%$ i łagodny poziom tolerancji — $p=30\%$, gdzie p oznacza procent brakujących informacji. Należy przy tym przestrzegać zasady, aby w jednym wierszu oraz jednej kolumnie macierzy X nie było więcej niż p informacji brakujących. Jeżeli procent braku danych jest większy niż p , należy z analizy wyłączyć odpowiedni obiekt lub określoną cechę. W sytuacji gdy nie zostaje przekroczona wartość p , możemy oszacować brakujące dane statystyczne na podstawie posiadanych informacji lub przeprowadzić klasyfikację na podstawie niekompletnych danych. Przy uzupełnianiu brakujących danych korzysta się z różnych procedur numerycznych, np. wykorzystuje się interpolację i ekstrapolację trendu, metodę najbliższego sąsiedztwa czy przeprowadza wnioskowanie na podstawie modeli ekonometrycznych.

Kolejnym etapem w wyodrębnianiu jednorodnych grup jest budowa macierzy podobieństwa obiektów, czyli macierzy odległości. Przez pojęcie odległości $d(O_i, O_j)$ między obiektami O_i i O_j należy rozumieć odległość między punktami (reprezentującymi te obiekty) w wielowymiarowej przestrzeni, która wyznaczana jest przez reguły określone w geometrii analitycznej. Wyznaczone odległości pozwalają określić położenie każdego punktu w stosunku do pozostałych punktów, a tym samym wskazać miejsce tego punktu w całej zbiorowości i zarazem umożliwić ich uporządkowanie i klasyfikację (Hellwig, Walter, 1979). **Odległości między obiektami** mogą być wyznaczone w różny sposób. Ich pod-

stawowy podział dotyczy sposobu konstrukcji tej miary. Pierwsza z nich (należą tu m.in. metody: najbliższego sąsiedztwa, najdalszego sąsiedztwa, mediany, średniej grupowej, środka ciężkości, metryka miejska oraz euklidesowa) opiera się na wyznaczeniu miary odległości jako bezwzględnej różnicy między wartościami cech diagnostycznych. Druga miara (dotyczy takich miar, jak np.: współczynnik dywergencji P. J. Clarka, odległość od pewnego wzorca J. Kolonko, odległość łukowa, propozycje Jeffrey, Braya czy Curtisa) wyznaczana jest na podstawie względnych różnic wartości cech diagnostycznych. Wybór miary odległości zależy od celów badania, możliwości oceny i interpretacji uzyskanych wyników oraz zakresu i jakości dostępnych danych statystycznych (Młodak, 2006).

Istotnym pojęciem w metodach taksonomicznych jest klasyfikacja³. W szerokim znaczeniu termin ten oznacza:

- *czynność* podziału zbioru obiektów na podzbiory według ustalonego kryterium, gdzie podział przeprowadzany jest na podstawie podobieństwa obiektów, a przez podobieństwo obiektów należy rozumieć wspólność (zbieżność) pewnych właściwości dwu lub więcej obiektów;
- *efekt* czynności podziału zbioru, którym są grupy obiektów podobnych;
- *decyzję*, do której klasy zbioru zaliczyć obiekt.

W sposób formalny możemy go opisać jako podział danego n -elementowego zbioru Ω , zawierającego obiekty badania $O_1, O_2, \dots, O_n$, które opisywane są przez k cech diagnostycznych $X_1, X_2, \dots, X_k$, na p podzbiorów $A_1, A_2, \dots, A_p$, gdzie $1 \leq p \leq n$ tak, aby spełnione były warunki (Pluta, 1986):

- 1) $A_1 \cup A_2 \cup \dots \cup A_p = \Omega$ (warunek zupełności),
- 2) $A_i \cap A_j \neq \emptyset$ ($i, j = 1, 2, \dots, p; i \neq j$) (warunek rozłączności),
- 3) $A_i \neq \emptyset$ ($i = 1, 2, \dots, p$) (warunek niepustego zbioru).

Po dokonaniu klasyfikacji obiektów, kolejnym krokiem w postępowaniu jest przeprowadzenie oceny podobieństwa obiektów należących do danego skupiska (grupy) ze względu na opisujące je cechy diagnostyczne. W tym celu wykorzystuje się miary homogeniczności, heterogeniczności skupień oraz poprawności skupień.

Pierwsza grupa tych miar, wyznaczana zazwyczaj przy wykorzystaniu miar tendencji centralnej, pokazuje skalę podobieństwa między badanymi obiektami. Im mniejsze wartości, tym mniejsze różnice w wartościach cech, a zatem większe podobieństwo obiektów. W drugim przypadku są to miary pokazujące odrębność skupień i im wyższa ich wartość, tym większy dystans między skupiskami. Do trzeciej grupy zaliczamy te miary, które są skonstruowane jako iloraz

³ Konieczność klasyfikacji wynika z kilku przesłanek: *metodologicznej* — sprowadzającej się do uzyskiwania bardziej jednorodnych obiektów analizy, w których łatwo wyodrębnić czynniki systematyczne oraz określić związki przyczynowo-skutkowe; *poznawczej* — polegającej na zredukowaniu dużej ilości informacji do kilku podstawowych kategorii, co upraszcza wnioskowanie; *ekonomicznej* — pozwalającej na zmniejszenie nakładów pracy poprzez ograniczenie rozważań do typowych faktów, zjawisk, obiektów przy stosunkowo niewielkim zniekształceniu wyników badania (Grabiński i in., 1991).

miar homogenicznych i heterogenicznych, pamiętając o jednorodności konstrukcji licznika i mianownika tej miary. Im mniejsza wartość tego wskaźnika, tym grupowanie obiektów bardziej poprawne (Młodak, 2006).

W literaturze przedmiotu można znaleźć wiele propozycji konstrukcji miar podobieństwa. Do najczęściej wykorzystywanych, ze względu na ich prostotę, należy zaliczyć:

- do oceny stopnia podobieństwa wewnątrzgrupowego — średnią odległość wewnątrzgrupową określoną wzorem:

$$d_{pp} = \frac{1}{n_p(n_p - 1)} \sum_{O_i \in A_p} \sum_{O_j \in A_p} d(O_i, O_j)$$

- do oceny stopnia podobieństwa międzygrupowego, gdzie p i q to odległości dla odpowiednich skupisk — miarę wyrażoną wzorem:

$$d_{pq} = \frac{1}{n_p n_q} \sum_{O_i \in A_p} \sum_{O_j \in A_q} d(O_i, O_j)$$

- do oceny poprawności klasyfikacji:

$$K = \frac{1}{p} \sum_{p=1} d_{pp} \quad \text{lub} \quad F = \frac{1}{p(p-1)} \sum_{p=1} \sum_{q>p} d_{pq}$$

(Podolec, Sokołowski, 1984; Nowak, 1985).

Obecnie omówimy szczegółowo te metody taksonomiczne, które najczęściej wykorzystywane są przez badaczy w różnego rodzaju analizach społeczno-ekonomicznych. Dla nich też przeprowadzimy analizę porównawczą uzyskanych wyników, w zależności od doboru zmiennych opisujących cel badania oraz od sposobu sprowadzania ich do porównywalności. Pierwsza z nich oparta jest na taksonomicznym mierniku rozwoju. Miernik ten służy głównie do liniowego uporządkowania obiektów ze względu na rozwój jakiegoś zjawiska.

Klasyfikacja zbioru obiektów sprowadza się do podziału obiektów według jednej zmiennej syntetycznej. Zmienna ta może zawierać w sobie informacje tylko o jednej cesze, ale może być również utworzona z większej ich liczby⁴. Punktem wyjścia w tej metodzie jest uporządkowanie obiektów według nierosnących (niemalejących) wartości syntetycznego miernika, a następnie podział analizowanego zbioru na 4 (lub więcej)⁵ grupy przy wykorzystaniu relacji mię-

⁴ Zmienna ta konstruowana bywa zazwyczaj na podstawie jednej cechy.

⁵ W literaturze przedmiotu podział zbioru obiektów na podzbiory dotyczył zwykle 4, 6 lub 8 klas. Ich liczba uzależniona była od liczebności zbioru wyjściowego. Im był on liczniejszy, tym liczba klas większa.

dzy średnią arytmetyczną a odchyleniem standardowym (lub q -krotnością odchylenia standardowego⁶):

- grupa I — $x_i < \bar{x} - q \cdot S(x)$,
- grupa II — $\bar{x} - q \cdot S(x) \leq x_i < \bar{x}$,
- grupa III — $\bar{x} \leq x_i < \bar{x} + q \cdot S(x)$,
- grupa IV — $x_i \geq \bar{x} + q \cdot S(x)$.

Zaletą tej metody jest to, że pozostawia swobodę badaczowi w sposobie konstrukcji miary syntetycznej oraz wyodrębnieniu liczby grup obiektów podobnych. Poprawność wniosków uzyskanych z tych analiz w dużym stopniu zależy zatem od trafności kryteriów przyjętych przez badacza.

W metodzie aglomeracyjnej dokonuje się łączenia obiektów w grupy, które opisywane są przez wiele zmiennych. Początkowo każdy obiekt traktowany jest jako odrębna grupa; tyle jest skupisk, ile obiektów. W drugim kroku dwa obiekty łączone są w jedno skupisko. W trzecim podejściu inny obiekt jest dodawany do już powstałego skupiska lub tworzy nowe samodzielne skupisko. W następnych krokach inne pojedyncze obiekty dodawane są do już utworzonych skupień lub tworzą nowe skupienie. W efekcie żadne utworzone skupisko nie może być podzielone, może jednak składać się z wielu innych skupisk niższego poziomu aglomeracji. Metoda ta nie pozostawia obiektów poza skupieniem, nie tworzy obiektów od niego odseparowanych. Proces klasyfikacji kończy się w momencie uzyskania jednej grupy obejmującej wszystkie obiekty zbioru.

Efektywny podział na skupiska uzyskuje się po zastosowaniu odpowiedniego, ustalonego arbitralnie, progu przerwania grupowania. Połączenia najbliższe tej wartości wyznaczają optymalne skupiska. Propozycje analitycznego wyznaczenia wartości progowej są bardzo różne. Można znaleźć je m.in. w pracach F. Wysokiego (1996) oraz R. Jabłońskiego i A. Robaszek (2000). Najczęściej wartość progowa ustalana jest subiektywnie przez badacza w taki sposób, aby liczba powstałych podgrup nie była zbyt liczna i aby zawierały one więcej niż jeden obiekt.

W zależności od techniki grupowania wyróżnia się kombinatoryczne i niekombinatoryczne procedury aglomeracyjne. Różnica między nimi polega na przyjętych kryteriach łączenia poszczególnych par skupień oraz na formule wyznaczania odległości nowego skupienia od pozostałych. W metodzie kombinatorycznej mamy następujące kroki (Nowak, 1990):

- 1) na podstawie macierzy odległości **D** szuka się pary skupień p i q (pary obiektów) ($p < q$) o najmniejszej odległości od siebie według zasady:

$$d_{pq} = \min_{ij} \{d_{ij}\} \quad (i, j = 1, \dots, n)$$

gdzie n jest liczbą istniejących grup (w pierwszym kroku n oznacza liczbę obiektów);

⁶ Przegląd literatury pozwolił zauważyć, że najczęściej przyjmowaną wartością q były takie liczby, jak: 0,5; 1,0; 1,5; 2,0 lub 2,5.

- 2) łączy się skupienia p i q w jedno skupienie, nadając mu numer p i usuwając jednocześnie skupienie o numerze q ;
- 3) wyznacza się odległość nowego skupienia od wszystkich pozostałych skupień, stosując formułę:

$$d_{pr} = a \cdot d_{pr} + a \cdot d_{gr} + b \cdot d_{pq} + c \cdot |d_{pr} - d_{gr}|$$

gdzie r przybiera wartości różne od p i q , natomiast parametry a , b , c wyznaczone są na podstawie metody Warda, gdzie: $a = \frac{n_i + n_p}{n_i + n_r}$, $b = -\frac{n_i}{n_i + n_r}$,

$c=0$;

- 4) powtarza się kroki 1—3, redukując stopniowo wymiar macierzy $\mathbf{D}$ o jeden do chwili, gdy wszystkie obiekty utworzą jedną grupę. W wyniku końcowym dostaje się n jednorodnych grup obiektów.

Zaletą omawianej metody jest jej prostota oraz możliwość przedstawienia wyników klasyfikacji w formie graficznej za pomocą drzewka połączeń (dendrogramu), które ilustruje kolejne połączenia grup coraz to wyższego rzędu wraz z miernikami charakteryzującymi odległości między poszczególnymi grupami. Podziału tego zazwyczaj proponuje się dokonywać w miejscach o najmniejszej gęstości połączeń, tzn. odcinając najdłuższe gałęzie drzewa.

ZASTOSOWANIE METOD TAKSONOMICZNYCH W ANALIZIE PRZESTRZENNEGO ZRÓŻNICOWANIA RYNKU PRACY

W tej części artykułu przedstawimy analizę przestrzennego zróżnicowania rynku pracy. Nie interesuje nas jednak ocena tego zjawiska, ale jedynie zmienne i metody wykorzystane w opisie. Dlatego rozważania nasze dotyczyć będą podziału administracyjnego z 1998 r., czyli na 49 województw (odpowiadają one aktualnemu podziałowi kraju na podregiony), które należy traktować jako przykładowe obiekty/jednostki w tym badaniu. Uznaliśmy, że obecnie obowiązujący podział administracyjny kraju na województwa lub powiaty nie jest korzystny z punktu widzenia prowadzonych tu rozważań. W pierwszym przypadku informacje są zbyt ogólne, a w drugim — zbyt szczegółowe.

W analizach przestrzennego zróżnicowania bezrobocia najczęściej bada się takie cechy, jak: stopa bezrobocia (y_1), osoby bezrobotne według wieku⁷, osoby bezrobotne według wykształcenia⁸, odsetek osób bezrobotnych pozostających w tej populacji powyżej 12 miesięcy (y_4), odsetek osób bezrobotnych biernych zawodowo przed zarejestrowaniem (y_5). Niektóre z tych cech mają wiele wariantów.

⁷ Najczęściej są to pięcioletnie grupy wieku określone na przedziale 20—60/65 lat w zależności od płci badanych.

⁸ Poziom wykształcenia został określony jako: wyższy, średni, zasadniczy i podstawowy (w tym również niepełny podstawowy).

tów, zbadano więc korelację między nimi. Stwierdzono, że w przypadku niektórych grup wieku oraz poziomów wykształcenia występuje między nimi istotna zależność. Wybrano zatem zmienne reprezentanty tych cech, czyli osoby powyżej 35. roku życia (y_2) oraz osoby z wykształceniem poniżej zasadniczego (y_3).

Syntetyczne miary zróżnicowania wyznaczone dla zmiennych według województw wskazują, że charakteryzują się one dużym rozproszeniem (współczynnik zmienności przyjmuje wartości wyższe niż 0,6). Zmienne te będziemy traktować jako destymulanty, gdyż wzrost ich wartości określał pogorszenie sytuacji na rynku pracy danego województwa w stosunku do pozostałych.

Jednowymiarowy opis przestrzennego zróżnicowania bezrobocia przeprowadzono osobno dla każdej z wymienionych zmiennych, traktując je jako syntetyczne mierniki rozwoju. Podziału województw na 4 jednorodne grupy dokonamy przy wykorzystaniu relacji między średnią arytmetyczną a q -krotnością odchylenia standardowego⁹.

W analizie wielowymiarowej zastosowano dwa podejścia. W pierwszym skorzystano z relacji między klasycznymi miarami statystycznymi, ale syntetyczny miernik rozwoju wyznaczony zostanie jako suma rang z pozycji zajmowanych przez województwo na liście rankingowej wyznaczonej dla każdej ze zmiennych osobno. W drugim podejściu opis sytuacji na rynku pracy przeprowadzono z wykorzystaniem metody aglomeracyjnej po raz pierwszy, gdy zmienne te podane zostaną procedurze standaryzacji, a drugi raz, gdy przejdą procedurę normalizacji.

W przeprowadzonych analizach jednowymiarowych przyporządkowano obiekty do określonych klas wartości tych zmiennych. Im numer klasy wyższy, tym ocena sytuacji na rynku pracy była gorsza. Dla każdej z analizowanych zmiennych osobno uzyskano różną liczbę obiektów wchodzących w skład grup podobnych oraz w zależności od przyjętej zmiennej do oceny sytuacji w danym województwie. Ocena ta może być uznana jako dobra lub bardzo zła (tabl. 1).

Spostrzeżenia te prowadzą do wniosku, że w ocenie sytuacji na regionalnych rynkach pracy należy uwzględnić z jednej strony tylko te zmienne, które w sposób istotny wpływają na jej zróżnicowanie i dotyczą istotnych zagadnień tego rynku (dobór zmiennych uzależniony jest od celu badania), ale z drugiej strony wydaje się, że analizy takie powinny być prowadzone przy uwzględnieniu wielu zmiennych łącznie. Wtedy bowiem uzyskujemy nieco bardziej ogólny jego obraz.

W analizie wielowymiarowej oceny sytuacji na regionalnych rynkach pracy zbudowano zmienną syntetyczną, która jest sumą rang pięciu cech. Do podziału województw na podobne grupy wykorzystano relację między średnią a odchyleniem standardowym. Podejście to pozwoliło na wyodrębnienie pięciu grup obiektów podobnych, ale liczba województw wchodzących w ich skład jest nieco odmienna. Ponadto, w ramach jednej grupy zawierają się województwa

⁹ Zróżnicowanie wartości tych zmiennych jest odmienne, nie można zatem zawsze przyjąć tych samych liczb dla q .

o dobrej i złej sytuacji na rynku pracy, ocenianej na podstawie tylko jednej z tych cech (tabl. 2).

TABL. 1. KLASYFIKACJA WOJEWÓDZTW WEDŁUG KLAS WARTOŚCI ZMIENNYCH OPISUJĄCYCH SYTUACJĘ NA RYNKU PRACY

Województwa	y ₁	y ₃	y ₄	y ₂	y ₅	Województwa	y ₁	y ₃	y ₄	y ₂	y ₅
Warszawskie	1	4	1	1	2	Olsztyńskie	4	1	3	2	3
Białkopodlaskie	2	4	3	4	3	Opolskie	2	2	2	2	2
Białostockie	2	4	1	3	3	Ostrołęckie	3	1	4	4	3
Bielskie	1	3	1	2	2	Piłskie	3	1	1	3	1
Bydgoskie	3	2	3	3	2	Piotrkowskie	3	3	4	2	3
Chełmskie	3	3	3	3	3	Płockie	3	3	3	3	4
Ciechanowskie	4	2	4	3	2	Poznańskie	1	4	1	2	2
Częstochowskie	2	2	1	2	1	Przemyskie	3	3	3	3	3
Elbląskie	4	1	2	2	1	Radomskie	3	1	4	3	2
Gdańskie	2	3	1	2	1	Rzeszowskie	3	2	2	4	3
Gorzowskie	3	2	1	2	2	Siedleckie	2	4	3	3	3
Jeleniogórskie	4	1	2	2	3	Sieradzkie	3	2	2	2	3
Kaliskie	3	2	3	3	2	Skierniewickie	2	1	3	2	4
Katowickie	1	3	1	1	4	Słupskie	4	2	2	2	1
Kieleckie	3	3	3	3	2	Suwałskie	4	2	3	2	1
Konińskie	3	2	4	3	4	Szczecińskie	2	4	1	1	2
Koszalińskie	4	2	2	2	1	Tarnobrzeskie	3	3	4	4	4
Krakowskie	1	4	1	3	2	Tarnowskie	2	2	3	4	4
Krośnińskie	3	1	4	4	3	Toruńskie	4	2	3	2	1
Legnickie	3	2	1	1	2	Wałbrzyskie	4	3	2	2	3
Leszczyńskie	2	2	2	2	2	Włocławskie	1	2	3	3	3
Lubelskie	2	4	3	3	3	Wrocławskie	1	3	1	1	3
Łomżyńskie	3	2	4	3	2	Zamojskie	2	3	4	4	4
Łódzkie	3	4	1	1	2	Zielonogórskie	3	2	1	2	2
Nowosądeckie	2	2	3	4	3						

Źródło: opracowanie własne.

TABL. 2. JEDNORODNE GRUPY WOJEWÓDZTW CHARAKTERYZOWANE PRZEZ ZMIENNĄ SYNTETYCZNĄ RANG W OPISIE SYTUACJI NA RYNKU PRACY

Grupy	Województwa
1	częstochowskie, warszawskie, bielskie, gdańskie, legnickie, piłskie, wrocławskie (n=7)
2	elbląskie, gorzowskie, katowickie, leszczyńskie, opolskie, poznańskie, szczecińskie, zielonogórskie, koszalińskie, krakowskie, łódzkie, słupskie, jeleniogórskie, sieradzkie, skierniewickie, suwałskie, toruńskie, włocławskie (n=18)
3	białostockie, bydgoskie, kaliskie, olsztyńskie, radomskie, kieleckie, łomżyńskie, nowosądeckie, rzeszowskie (n=9)
4	wałbrzyskie, chełmskie, ciechanowskie, krośnińskie, lubelskie, ostrołęckie, piotrkowskie, przemyskie, siedleckie, tarnowskie, białkopodlaskie, konińskie, płockie (n=13)
5	zamojskie, tarnobrzeskie (n=2)

Źródło: jak przy tabl. 1.

W naszej analizie zastosowano również metodę aglomeracyjną. Wszystkie wykorzystywane zmienne spełniały kryterium poziomu i rozproszenia oraz braku korelacji. Każdą ze zmiennych traktowano jako tak samo istotną. Ich wagi równe są jedności. Równocześnie założono, że w pierwszej z rozważanych sytuacji każda ze zmiennych wywiera jednakowy wpływ na odległość między badanymi obiektami, zaś w drugiej sytuacji wpływ ten nie jest jednakowy. Zmienne podlegają zatem procedurze standaryzacji lub normalizacji. Kolejnym etapem analizy jest wyodrębnianie jednorodnych obszarów na podstawie macierzy odległości metodą Warda. Wyniki analizy zawierają tabl. 3 i 4.

TABL. 3. JEDNORODNE GRUPY WOJEWÓDZTW CHARAKTERYZOWANE PRZEZ ZMIENNE STANDARYZOWANE W OPISIE SYTUACJI NA RYNKU PRACY

Grupy ^a	Województwa
1	elbąskie, gorzowskie, jeleniogórskie, koszańskie, legnickie, olsztyńskie, piotrkowskie, radomskie, śląskie, suwalskie, toruńskie, wałbrzyskie, wrocławskie (n=13)
2	białopodlaskie, bydgoskie, ciechanowskie, chełmskie, kaliskie, konińskie, krośnieńskie, łomżyńskie, ostrołęckie, nowosądeckie, płockie, przemyskie, siedleckie, sieradzkie, skierniewickie, rzeszowskie, tarnobrzyskie, tarnowskie, zamojskie (n=19)
3	białostockie, bielskie, częstochowskie, gdańskie, leszczyńskie, katowickie, kieleckie, lubelskie, opolskie, poznańskie, pilskie, szczecińskie, wrocławskie, zielonogórskie (n=14)
4	warszawskie (n=1)
5	krakowskie (n=1)
6	łódzkie (n=1)

^a Grupy województw przedstawiono według rosnącej wartości odległości.

Źródło: jak przy tabl. 1.

TABL. 4. JEDNORODNE GRUPY WOJEWÓDZTW CHARAKTERYZOWANE PRZEZ ZMIENNE ZNORMALIZOWANE W OPISIE SYTUACJI NA RYNKU PRACY

Grupy ^a	Województwa
1	bydgoskie, wałbrzyskie, elbąskie, śląskie, gorzowskie, jeleniogórskie, wrocławskie, koszańskie, legnickie, łomżyńskie, ostrołęckie, pilskie, piotrkowskie, toruńskie, skierniewickie, zielonogórskie (n=16)
2	białostockie, płockie, kieleckie, konińskie, łódzkie, ciechanowskie, olsztyńskie, radomskie, suwalskie (n=9)
3	częstochowskie, siedleckie, kaliskie, białopodlaskie, chełmskie, krośnieńskie, leszczyńskie, tarnowskie, lubelskie, nowosądeckie, przemyskie, rzeszowskie, sieradzkie, tarnobrzyskie, zamojskie (n=15)
4	bielskie, opolskie, gdańskie, poznańskie, szczecińskie, katowickie, wrocławskie (n=7)
5	warszawskie, krakowskie (n=2)

^a Jak przy tabl. 3.

Źródło: jak przy tabl. 1.

Dla zmiennych poddanych procedurze standaryzacji wyodrębniono 6 jednorodnych grup województw, charakteryzujących się różną liczbą obiektów wchodzących w ich skład (trzy z nich zawierają po jednym elemencie). W przypadku zmiennych poddanych procedurze normalizacji uzyskano 5 grup województw, również o różnej liczbie obiektów wchodzących w ich skład, ale bez skupisk jednoelementowych. Porównując wyniki uzyskane dla tych zgrupowań można zauważyć, że w skład wyodrębnionych grup nie wchodzi te same województwa. Niektóre z nich tworzą skupienia z innymi obiektami lub pozostają grupami jednoelementowymi. W tabl. 5 przedstawiono obiekty tworzące podobne skupiska dla pierwszej i drugiej sytuacji. Do podobnych wniosków dochodzi się, porównując utworzone grupy województw podobnych oraz elementy wchodzące do grupy z wynikami uzyskanymi w analizie wykorzystującej syntetyczny miernik rozwoju jako sumę rang.

TABL. 5. GRUPY WOJEWÓDZTW TWORZĄCE PODOBNE SKUPISKA DLA ZMIENNYCH STANDARYZOWANYCH I ZNORMALIZOWANYCH W OPISIE SYTUACJI NA RYNKU PRACY

Grupy	Województwa
1	kaliskie, chełmskie, sieradzkie, przemyskie, rzeszowskie, białkopodlaskie, siedleckie, nowosądeckie, zamojskie, krośnieńskie, tarnobrzskie, tarnowskie ($n=12$)
2	łomżyńskie, suwalskie, skierniewickie, ostrołęckie, olsztyńskie, radomskie, ciechanowskie, bydgoskie, płockie, konińskie ($n=10$)
3	piotrkowskie, koszalińskie, słupskie, elbląskie, gorzowskie, legnickie, wałbrzyskie, wrocławskie, jeleniogórskie, toruńskie ($n=10$)
4	bielskie, katowickie, gdańskie, opolskie, wrocławskie, szczecińskie, poznańskie ($n=7$)
5	białostockie, częstochowskie, leszczyńskie, lubelskie, kieleckie ($n=5$)
6	pilskie, zielonogórskie ($n=2$)
7	łódzkie ($n=1$)
8	warszawskie ($n=1$)
9	krakowskie ($n=1$)

Źródło: obliczenia własne na podstawie tabl. 3 i 4.

Istotnym elementem przy wyodrębnianiu homogenicznych grup jest ocena dokonanego podziału przestrzeni wielowymiarowej na jednorodne podprzestrzenie. W tym celu wykorzystuje się miary stopnia podobieństwa wewnątrzgrupowego oraz międzygrupowego. W metodzie aglomeracyjnej dla obu rozważanych sytuacji uzyskano relatywnie duże podobieństwo obiektów należących do danego skupiska oraz względnie małe podobieństwo między nimi. Średnia odległość wewnątrzgrupowa dla zmiennych standaryzowanych wynosiła 0,9, a zmiennych normalizowanych — 2,2. Średnia odległość międzygrupowa była odpowiednio prawie 8-krotnie i 9-krotnie wyższa. Wyznaczając miarę popraw-

ności skupień jako iloraz odległości wewnątrzgrupowej i międzygrupowej uzyskano wartości 0,11 i 0,24. Wskazuje to, że jednorodne grupy województw wyodrębnione ze względu na ocenę sytuacji na rynku pracy były poprawne.

Reasumując w wielowymiarowej analizie statystycznej, przy uwzględnieniu tych samych zmiennych, ale różnego sposobu sprowadzania ich do porównywalności, uzyskano 5 lub 6 grup województw podobnych, o różnej liczbie obiektów wchodzących w ich skład. Ponadto, w skład tych grup nie zawsze wchodziły te same obiekty. Porównując województwa wchodzące w skład tych samych skupień z ich przynależnością do klas utworzonych w analizie jednowymiarowej można zauważyć, że tylko w niewielkim stopniu się pokrywają. Nie oznacza to jednak, że ocena sytuacji na rynku pracy przy wykorzystaniu rozmaitej liczby zmiennych oraz różnych metod wyodrębniania grup obiektów podobnych może prowadzić do błędnych wyników. Każda z tych sytuacji opisuje zjawisko z innej strony. Dlatego tak ważne jest precyzyjne ustalenie celu badania, określenie własności zmiennych w nich wykorzystywanych oraz dobór odpowiedniej metody.

Uwagi końcowe

Popularność stosowania metod taksonomicznych sugeruje uniwersalną poprawność ich stosowania. Niestety, nie zawsze się tak dzieje. Często do opisu badanego zjawiska wykorzystuje się wszystkie znane (możliwe do wykorzystania) cechy, nie zastanawiając się czy można wykazać ich logiczny związek z opisywaną sytuacją albo czy zmienne wykorzystywane w analizie nie powielają informacji. Ponadto nie przywiązuje się wagi do przekształcania zmiennych w ten sposób, aby ich skala oraz kierunek wpływu na opisywane zjawisko były zgodne z innymi zmiennymi. Dobór metody do analizy jest też nierzadko przypadkowy, nieuwzględniający właściwości zmiennych w niej wykorzystywanych. Wszystko to powoduje, że wnioski uzyskane z prowadzonych analiz mogą być obciążone dużymi błędami.

dr Małgorzata Podogrodzka — SGH

LITERATURA

- Bartosiewicz S. (1976), *Wariacje na temat wzorców rozwoju*, „Przegląd Statystyczny”, nr 3, *Encyklopedia Popularna* (1982), PWN, Warszawa
- Gorzelak G. (1981), *Statystyczna analiza porównawcza — teoria a praktyka*, „Wiadomości Statystyczne”, nr 8
- Grabiński T., Ludwiczak B., Malina A., Zeliaś A. (1991), *Ekonometria przestrzenna*, PWE, Warszawa
- Grabiński T., Szymanowicz K., Woźniak M., Zeliaś A. (1976), *O pewnej metodzie grupowania zmiennych*, „Przegląd Statystyczny”, nr 4

- Hellwig Z. (1968), *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr*, „Przegląd Statystyczny”, nr 4
- Hellwig Z., Walter C. (1979), *Taksonomiczne i ekonometryczne podejście do analizy danych statystycznych*, „Wiadomości Statystyczne”, nr 9
- Jabłoński R., Robaszek A. (2000), *Przestrzenne zróżnicowanie warunków środowiskowych w powiatach województwa łódzkiego*, Seria: Analizy Statystyczne, Urząd Statystyczny w Łodzi
- Kowalski A. (1977), *Charakterystyka niektórych metod taksonomicznych*, PWN, Warszawa
- Młodak A. (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa
- Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa
- Nowak E. (1985), *Metoda statystycznych analiz porównawczych efektywności obiektów rolniczych*, „Prace Naukowe Akademii Ekonomicznej we Wrocławiu”, nr 292
- Pluta W. (1986), *Wielowymiarowa analiza porównawcza w badaniach ekonometrycznych*, PWN, Warszawa
- Pluta W. (1977), *Wielowymiarowa analiza porównawcza w badaniach ekonomicznych. Metody taksonomiczne i analizy czynnikowej*, PWE, Warszawa
- Podolec B., Sokołowski A. (1984), *Metody badania przestrzennej zwartości podzbiorów otrzymanych w wyniku taksonomicznego grupowania jednostek terytorialnych*, „Prace Naukowe Akademii Ekonomicznej we Wrocławiu”, nr 262
- Strahl D. (1980), *Modelowanie zjawisk złożonych. Modele infrastruktury społecznej*, „Prace Naukowe Akademii Ekonomicznej we Wrocławiu”, nr 158
- Styczeń M. (1971), *Ogólna charakterystyka metod taksonomicznych*, „Wiadomości Statystyczne”, nr 8
- Wysocki F. (1996), *Metody statystycznej analizy wielowymiarowej w rozpoznawaniu typów struktury przestrzennej rolnictwa*, Roczniki Akademii Rolniczej im. A. Cieszkowskiego w Poznaniu, Poznań

SUMMARY

Popularity of methods used in the taxonomic analysis of comparative socio-economic phenomena contributes to reflect on the correctness of their application. The article presents the research topics that use these methods and describes the proper conditions for their use, with particular emphasis on methods of agglomeration and synthetic development measure.

РЕЗЮМЕ

Популярность использования таксономических методов в сопоставительном анализе социально-экономических явлений является причиной рассмотрения правильности их использования. В статье были представлены обследовательские темы использующие эти методы, а также были охарактеризованы условия правильного их использования, обращая особое внимание на агломерационный метод и метод синтетического измерителя развития.