

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 328

Taksonomia 23

**Klasyfikacja i analiza danych –
teoria i zastosowania**

Redaktorzy naukowci

Krzysztof Jajuga, Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2014

Redaktor Wydawnictwa: Barbara Majewska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

www.ibuk.pl, www.ebscohost.com,

w Dolnośląskiej Bibliotece Cyfrowej www.dbc.wroc.pl,

The Central and Eastern European Online Library www.ceeol.com,

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się
na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego
oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2014

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp	11
Małgorzata Rószkiewicz , Wykorzystanie metaanalizy w budowaniu modelu pomiarowego w przypadku braku niezmienniczości zasad pomiaru na przykładzie pomiaru zadowolenia z życia.....	13
Elżbieta Sobczak , Harmonijność inteligentnego rozwoju regionów Unii Europejskiej	21
Ewa Roszkowska, Renata Karwowska , Analiza porównawcza województw Polski ze względu na poziom zrównoważonego rozwoju w roku 2010.....	30
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel , Analiza porównawcza wybranych filtrów w analizie synchronizacji cyklu koniunkturalnego.....	41
Marcin Salamaga , Próba konstrukcji tablic „wymierania scenicznego” spektakli operowych na przykładzie Metropolitan Opera.....	51
Iwona Foryś , Wykorzystanie analizy dyskryminacyjnej do typowania rynków podobnych w procesie wyceny nieruchomości niemieszkalnych	59
Jerzy Korzeniewski , Selekcja zmiennych w klasyfikacji – propozycja algorytmu	69
Sabina Denkowska , Testowanie wielokrotne przy weryfikacji wieloczynnikowych modeli proporcjonalnego hazardu Coxa.....	76
Ewa Chodakowska , Teoria równań strukturalnych w klasyfikacji zmiennych jawnych i ukrytych według charakteru ich wzajemnych oddziaływań	85
Iwona Konarzewska , Model PCA dla rynku akcji – studium przypadku	94
Katarzyna Wójcik, Janusz Tuchowski , Dobór optymalnego zestawu słów istotnych w opiniach konsumentów na potrzeby ich automatycznej analizy	106
Aleksandra Łuczak , Zastosowanie metody AHP-LP do oceny ważności determinant rozwoju społeczno-gospodarczego w jednostkach administracyjnych	116
Aleksandra Witkowska, Marek Witkowski , Klasyfikacja pozycyjna banków spółdzielczych według stanu ich kondycji finansowej w ujęciu dynamicznym	126
Adam Depta , Zastosowanie analizy korespondencji do oceny jakości życia ludności na podstawie kwestionariusza SF-36v2	135
Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Marek Marciniak, Jerzy Kołodziej , Indukcja reguł dla danych niekompletnych i niezbalansowanych: modele klasyfikatorów i próba ich zastosowania do predykcji ryzyka operacyjnego w torakochirurgii	146

Małgorzata Misztal , Wybrane metody oceny jakości klasyfikatorów – przegląd i przykłady zastosowań.....	156
Anna M. Olszewska , Wykorzystanie wybranych metod taksonomicznych do oceny potencjału innowacyjnego województw	167
Iwona Bąk , Porównanie jakości grupowań powiatów województwa zachodniopomorskiego pod względem atrakcyjności turystycznej.....	177
Agnieszka Kozera, Joanna Stanisławska, Romana Głowicka-Wołoszyn , Segmentacja gospodarstw domowych według wydatków na turystykę zorganizowaną.....	186
Agnieszka Wałęga , Podejście syntetyczne w analizie spójności ekonomicznej gospodarstw domowych.....	196
Joanna Banaś, Małgorzata Machowska-Szewczyk, Bożena Mroczek , Zastosowanie analizy korespondencji do badania wpływu elektrowni wiatrowych na jakość życia ludności	205
Joanna Banaś, Krzysztof Małecki , Klasyfikacja punktów pomiarów ankietowych kierowców na granicy Szczecina z wykorzystaniem zmiennych symbolicznych.....	214
Aneta Becker , Wykorzystanie informacji granularnej w analizie wymagań rynku pracy.....	222
Katarzyna Cheba, Joanna Holub-Iwan , Wykorzystanie analizy korespondencji w segmentacji rynku usług medycznych.....	230
Adam Depta, Iwona Staniec , Identyfikacja czynników decydujących o jakości życia studentów łódzkich uczelni.....	238
Katarzyna Dębowska, Jarosław Kilon , Reguły asocjacyjne w analizie wyników badań metodą Delphi.....	247
Anna Domagała , O wykorzystaniu analizy głównych składowych w metodzie <i>Data Envelopment Analysis</i>	254
Alicja Grześkowiak , Analiza wykluczenia cyfrowego w Polsce w ujęciu indywidualnym i regionalnym.....	264
Anna M. Olszewska, Anna Gryko-Nikitin , Pomiar postrzegania jakości kształcenia uczelni wyższej na danych porządkowych z wykorzystaniem środowiska R.....	273
Karolina Paradysz , Hierarchiczna metoda grupowania powiatów jako podejście benchmarkowe w ocenie bezrobocia według BAEL-u w wybranych typach małych obszarów	282
Radosław Pietrzyk , Porównanie metod pomiaru efektywności zarządzania portfelami funduszy inwestycyjnych.....	290
Agnieszka Przedborska, Małgorzata Misztal , Wybrane metody statystyki wielowymiarowej w ocenie skuteczności terapeutycznej głębokiej stymulacji elektromagnetycznej u pacjentów z chorobą zwyrodnieniową stawów.....	299

Wojciech Roszka, Marcin Szymkowiak , Podejście kalibracyjne w statystycznej integracji danych	308
Iwona Skrodzka , Zastosowanie wybranych metod klasyfikacji do analizy kapitału ludzkiego krajów Unii Europejskiej	316
Agnieszka Stanimir , Wielowymiarowa analiza czynników sprzyjających włączeniu społecznemu	326
Dorota Strózik, Tomasz Strózik , Przestrzenne zróżnicowanie poziomu życia w województwie wielkopolskim.....	334
Izabela Szamrej-Baran , Identyfikacja przyczyn ubóstwa energetycznego w Polsce przy wykorzystaniu modelowania miękkiego.....	343
Janusz Tuchowski, Katarzyna Wójcik , Klasyfikacja obiektów w systemie Krajowych Ram Kwalifikacji opisanych za pomocą ontologii	353
Aleksandra Matuszewska-Janica , Grupowanie krajów Unii Europejskiej ze względu na poziom feminizacji sektorów gospodarczych	361
Monika Rozkrut, Dominik Rozkrut , Identyfikacja strategii innowacyjnych przedsiębiorstw usługowych w Polsce	369

Summaries

Małgorzata Rószkiewicz , The use of meta-analysis in building the measurement model in case of the absence of measurement invariance on the example of measuring of life satisfaction.....	20
Elżbieta Sobczak , Harmonious smart growth of European Union regions.....	29
Ewa Roszkowska, Renata Karwowska , The comparative analysis of Polish voivodeships with respect to sustainable development in 2010.....	40
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel , Comparative analysis of chosen filters in business cycles analysis	50
Marcin Salamaga , The attempt of construction of the life tables for opera works on the example of the Metropolitan Opera	58
Iwona Foryś , Using discriminant analysis to select similar markets in non-residential property valuation process.....	68
Jerzy Korzeniewski , Variable selection in classification – algorithm proposal	75
Sabina Denkowska , Multiple testing in the verification process of multifactorial Cox proportional hazards models	84
Ewa Chodakowska , The theory of structural equations modelling in the classification of observed variables and latent constructs according to the character of their relationship.....	93
Iwona Konarzewska , Modelling stock market by PCA factor model – case study	105

Katarzyna Wójcik, Janusz Tuchowski , Selection of the optimal set of relevant words in consumers opinions in the context of the opinion mining ..	115
Aleksandra Łuczak , Application of AHP-LP to the evaluation of importance of determinants of socio-economic development in the administrative units	125
Aleksandra Witkowska, Marek Witkowski , A dynamic approach to the ranking of cooperative banks by their financial condition	134
Adam Depta , Application of correspondence analysis for the measurement of quality of life – questionnaire SF-36v2 based research	145
Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Marek Marciniak, Jerzy Kołodziej , Classification rules extraction for missing and imbalance data: models of classifiers and initial results in the rules-based thoracic surgery risk prediction.....	155
Małgorzata Misztal , Selected methods for assessing the performance of classifiers – an overview and examples of applications.....	166
Anna M. Olszewska , The application of selected quantitative methods to the evaluation of voivodeship innovation level potential.....	176
Iwona Bąk , The comparison of the quality of groupings of poviats of West Pomeranian Voivodeship in terms of tourism attractiveness	185
Agnieszka Kozera, Joanna Stanisławska, Romana Głowicka-Wołoszyn , Household segmentation with respect to the expenditure on organized tourism.....	195
Agnieszka Wałęga , Synthetic approach in the analysis of economic coherence of households	204
Joanna Banaś, Małgorzata Machowska-Szewczyk, Bożena Mroczek , Using the correspondence analysis to examine the impact of wind turbines on the quality of life.....	213
Joanna Banaś, Krzysztof Małecki , Classification of measurement survey points of drivers on the boundary of Szczecin using symbolic variables...	221
Aneta Becker , The use granular information in the analysis of the requirements of the labor market.....	229
Katarzyna Cheba, Joanna Hołub-Iwan , The application of the correspondence analysis of patients segmentation on the medical service market	237
Adam Depta, Iwona Staniec , Identification of the factors that determine the quality of students life at universities in Lodz.....	246
Katarzyna Dębkowska, Jarosław Kilon , Association rules in the analysis of research results the Delphi method	253
Anna Domagała , About using Principal Component Analysis in Data Envelopment Analysis	263
Alicja Grześkowiak , Analysis of the digital divide in Poland at the individual and regional level	272

Anna M. Olszewska, Anna Gryko-Nikitin , Assessment of perception of quality of teaching at an institution of higher learning based on the ordinal data with the utilization of R environment.....	281
Karolina Paradysz , The hierarchical method of grouping poviats as a benchmark approach in the assessment of unemployment by BAEL in selected types of small areas	289
Radosław Pietrzyk , Comparison of methods of measuring the performance of investment funds portfolios.....	298
Agnieszka Przedborska, Małgorzata Misztal , Selected multivariate statistical analysis methods in the evaluation of efficacy of deep electromagnetic stimulation in patients with degenerative joint disease	307
Wojciech Roszka, Marcin Szymkowiak , A calibration approach in statistical data integration	315
Iwona Skrodzka , Application of some methods of classification to the analysis of human capital in the European Union.....	325
Agnieszka Stanimir , Multivariate analysis of social inclusion factors.....	333
Dorota Strózik, Tomasz Strózik , Spatial differentiation of the standard of living in Great Poland Voivodeship	342
Izabela Szamrej-Baran , Identification of fuel poverty causes in Poland using soft modelling	352
Janusz Tuchowski, Katarzyna Wójcik , Classification of objects in the National Classification Framework described by the ontology.....	360
Aleksandra Matuszewska-Janica , Clustering of European Union states taking into consideration the levels of feminization of economic sectors..	368
Monika Rozkrut, Dominik Rozkrut , Identification of service sector innovation strategies in Poland.....	379

Iwona Bąk

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

PORÓWNANIE JAKOŚCI GRUPOWAŃ POWIATÓW WOJEWÓDZTWA ZACHODNIOPOMORSKIEGO POD WZGLĘDEM ATRAKCYJNOŚCI TURYSTYCZNEJ

Streszczenie: Celem artykułu jest porównanie jakości grupowań powiatów województwa zachodniopomorskiego pod względem ich atrakcyjności turystycznej przy wykorzystaniu różnych metod doboru cech diagnostycznych (metoda parametryczna, metoda odwróconej macierzy) oraz klasycznych i pozycyjnych taksonomicznych mierników rozwoju. W analizie wykorzystano cechy diagnostyczne charakteryzujące walory przyrodnicze, stopień zagospodarowania turystycznego, dostępność komunikacyjną oraz poziom zanieczyszczenia środowiska naturalnego. Jakość grupowań sprawdzono, obliczając wskaźniki homogeniczności, heterogeniczności i poprawności skupień.

Słowa kluczowe: metoda doboru cech, taksonomiczny miernik rozwoju Hellwiga, mediana Webera, jakość grupowań, atrakcyjność turystyczna.

1. Wstęp

Coraz większa swoboda przepływu ludzi, towarów i kapitału, przy wyraźnym bogaceniu się społeczeństw, sprzyja rozwojowi turystyki. Z jej rozwojem wiąże się także wzrost konkurencji w turystyce nie tylko wśród przedsiębiorstw turystycznych, lecz również dotyczący miejscowości i regionów recepcji ruchu turystycznego. Zwiększenie konkurencyjności obszarów recepcji, m.in. poprzez rozwój turystyki i niezbędnej infrastruktury, można uznać za istotny czynnik bogacenia się jednostek samorządu terytorialnego i lokalnych społeczności [Rapacz (red.) 2012]. Na natężenie ruchu turystycznego wpływa przede wszystkim atrakcyjność turystyczna danego regionu, na którą składają się: walory przyrodnicze, stopień zagospodarowania turystycznego, dostępność komunikacyjna oraz poziom zanieczyszczenia środowiska naturalnego. Kompleksowe występowanie wymienionych czynników w pełni zaspokaja zapotrzebowanie turystów, a obszar, miejscowość czy szlak stają się atrakcyjne turystycznie.

Celem artykułu jest porównanie jakości grupowań powiatów województwa zachodniopomorskiego pod względem ich atrakcyjności turystycznej przy wykorzystaniu:

- różnych metod doboru cech diagnostycznych (metoda parametryczna, metoda odwróconej macierzy),
 - klasycznych i pozycyjnych taksonomicznych mierników rozwoju.
- Jakość grupowań sprawdzono, obliczając wskaźniki homogeniczności, heterogeniczności i poprawności skupień.

2. Metoda i materiał badawczy

W każdej wielowymiarowej analizie porównawczej na wstępie badania konieczne jest określenie obiektów porównania oraz zestawu cech w sposób wszechstronny charakteryzujących własności tych obiektów, czyli tzw. cech diagnostycznych. Istnieją dwa podstawowe podejścia do wyboru cech: merytoryczne i statystyczne [Wysocki 2010]. W podejściu pierwszym do zbioru cech diagnostycznych powołuje się takie wielkości, które w świetle posiadanej wiedzy merytorycznej o badanych zjawiskach stanowią najważniejsze charakterystyki porównywanych obiektów. W podejściu drugim wybór cech odbywa się w drodze przetwarzania i analizy informacji statystycznych za pomocą odpowiednich procedur statystycznych ze względu na następujące podstawowe kryteria: zdolność dyskryminacyjną cech, czyli ich zmienność względem badanych obiektów, oraz potencjał informacyjny cech, czyli stopień ich skorelowania z innymi cechami [Panek 2009, s. 19].

W artykule do badania atrakcyjności powiatów województwa zachodniopomorskiego wykorzystano dwie metody doboru cech diagnostycznych: metodę parametryczną, zaproponowaną przez Z. Hellwiga oraz metodę odwróconej macierzy. Otrzymane za pomocą tych metod zbiory cech stanowiły podstawę klasyfikacji powiatów, której dokonano za pomocą zmiennej syntetycznej. Syntetyczne mierniki rozwojowe umożliwiają uporządkowanie i wyznaczenie grup obiektów podobnych pod względem badanego zjawiska. W tym celu wykorzystuje się mierniki wzorcowe i bezwzorcowe [Grabiński 1992, s. 25]. W opracowaniu do budowy taksonomicznego miernika rozwoju zastosowano metodę wzorcową, w podejściu zarówno klasycznym, jak i pozycyjnym wykorzystującym medianę Webera. Budowa taksonomicznego miernika rozwoju w podejściu klasycznym omówiona została np. w pracy [Nowak 1990, s. 87-91].

Wariant pozycyjny przyjmuje odmienną formułę standaryzacyjną w porównaniu z podejściem klasycznym, opartą na ilorazie odchylenia wartości cechy od odpowiedniej współrzędnej mediany Webera i ważonego medianowego odchylenia bezwzględne z zastosowaniem mediany Webera [Młodak 2006, s. 136-137]:

$$z_{ij} = \frac{x_{ij} - \theta_{0j}}{1,4826 \cdot \text{m}\ddot{a}\text{d}(X_j)},$$

gdzie: $\theta_0 = (\theta_{01}, \theta_{02}, \dots, \theta_{0m})$ jest medianą Webera, $\text{m}\ddot{a}\text{d}(X_j)$ to medianowe odchylenie bezwzględne, w którym bada się dystans cech do wektora Webera, tzn.:

$$\text{m}\ddot{a}\text{d}(X_j) = \text{med}_{i=1,2,\dots,n} |x_{ij} - \theta_{0j}| \quad (j = 1, 2, \dots, m).$$

Miernik agregatowy wyznacza się według wzoru:

$$\mu_i = 1 - \frac{d_i}{d_-},$$

gdzie: $d_- = \text{med}(\mathbf{d}) + 2,5\text{mad}(\mathbf{d})$,

$\text{med}(\mathbf{d})$ – mediana z wektora odległości \mathbf{d} ,

$\text{mad}(\mathbf{d})$ – medianowe odchylenie bezwzględne wyznaczone jako funkcja modułów odchylenia odległości od ich mediany, przy czym $\mathbf{d} = (d_1, d_2, \dots, d_n)$ jest wektorem odległości wyznaczanym według wzoru:

$$d_i = \text{med}_{j=1,2,\dots,m} |z_{ij} - \phi_j| \quad i = 1, 2, \dots, n,$$

$\phi_j = \max_{i=1,2,\dots,n} z_{ij}$ – współrzędne wektora wzorca rozwoju, którymi są maksymalne wartości znormalizowanych cech.

Uporządkowania obiektów, zarówno za pomocą miernika klasycznego, jak i pozycyjnego stanowią podstawę do grupowania obiektów na cztery klasy. Stosowana najczęściej metoda grupowania w ujęciu klasycznym nosi nazwę metody trzech średnich [Nowak 1990, s. 93], a w ujęciu pozycyjnym trzech median [Młodak 2006, s. 138].

Do oceny jakości analizy taksonomicznej wykorzystano mierniki homogeniczności, heterogeniczności oraz poprawności skupień, oparte na koncepcji środka ciężkości grupy i odległości od niej. W pracy zastosowano podejście, w którym środek ciężkości danej grupy zastąpiony został medianą Webera jej elementów. W ocenie homogeniczności otrzymanych grup wykorzystano następujący miernik [Młodak 2006, s. 138-141]:

$$hm_6^* mx = \max_{k=1,2,\dots,p} hm_6^*(P_k),$$

gdzie: $hm_6^*(P_k) = \text{med}_{i:\Gamma_i \in P_k} \delta(\Gamma_i, \Gamma_{\theta k})$ to mediana odległości obiektów grupy P_k od jej wektora medianowego Webera,

$\Gamma_{\theta k} = (\theta_{1P_k}, \theta_{2P_k}, \dots, \theta_{mP_k})$ – wektor medianowy Webera wyznaczony dla grupy P_k , $k = 1, 2, \dots, p$,

p – liczba skupień otrzymanych na danym poziomie grupowania.

Natomiast w ocenie heterogeniczności zastosowano miernik:

$$ht_6^* mn = \min_{k=1,2,\dots,p} ht_6^*(P_k),$$

gdzie: $ht_6^*(P_k) = \text{med}_{\substack{i=1,\dots,p \\ i \neq k}} \delta(\Gamma_i, \Gamma_{\theta k})$ – mediana odległości pomiędzy medianą Webera

danej grupy z analogicznymi wektorami dla pozostałych grup.

W ocenie poprawności grupowania wykorzystano kompleksowy miernik o postaci:

$$ct_6^* = \frac{hm_6^*mx}{ht_6^*mn}.$$

Do opisu atrakcyjności turystycznej powiatów przyjęto następujący zestaw cech [US 2011]: X_1 – powierzchnia o szczególnych walorach przyrodniczych prawnie chroniona w % powierzchni ogólnej, X_2 – pomniki przyrody na 1 km², X_3 – lesistość w %, X_4 – ścieki przemysłowe i komunalne w dam³ na 1 km², X_5 – ludność korzystająca z oczyszczalni ścieków w % ludności ogółem, X_6 – emisja zanieczyszczeń pyłowych powietrza w tonach na 1 km², X_7 – emisja zanieczyszczeń gazowych powietrza w tonach na 1 km², X_8 – odpady wytworzone w ciągu roku ogółem w tys. ton na 1 km², X_9 – nakłady na środki trwałe służące ochronie środowiska w tys. zł, X_{10} – mieszkania na 1000 mieszkańców, X_{11} – drogi publiczne powiatowe o twardej nawierzchni w km na 100 km², X_{12} – targowiska stałe na 1000 ludności, X_{13} – stałe punkty sprzedaży drobnodetalicznej na 1000 ludności, X_{14} – podmioty gospodarki narodowej zarejestrowane w rejestrze REGON na 10 tys. ludności, X_{15} – łóżka w szpitalach na 10 tys. ludności, X_{16} – zakłady opieki zdrowotnej na 1000 ludności, X_{17} – liczba ludności na 1 aptekę, X_{18} – miejsca na widowni w kinach na 1000 ludności, X_{19} – obiekty turystyczne na 1000 ludności, X_{20} – miejsca noclegowe na 1000 ludności, X_{21} – liczba placówek bibliotecznych na 1000 ludności, X_{22} – muzea na 1000 ludności.

3. Wyniki badania

W pierwszym kroku uwzględniono kryterium dyskryminacyjne obiektów, wyrażone za pomocą współczynnika zmienności. Z uwagi na to, że współczynniki te przekraczały 10%, do badania przyjęto wszystkie zaproponowane cechy diagnostyczne¹. Po wyznaczeniu macierzy współczynników korelacji wyznaczono zbiory cech diagnostycznych, wykorzystując dwie metody ich doboru: metodę parametryczną oraz metodę odwróconej macierzy. Metoda parametryczna pozwoliła na wyodrębnienie następującego zbioru cech diagnostycznych: X_2 , X_3 , X_4 , X_{12} , X_{13} , X_{14} , X_{15} , X_{16} , X_{18} , X_{19} . Natomiast wykorzystanie metody odwróconej macierzy

¹ Z analizy cech diagnostycznych wynika, że badane powiaty znacznie różniły się między sobą ze względu na przyjęte cechy, o czym świadczą wysokie wartości współczynnika zmienności (od 13% do 340%). Ponadto większość cech charakteryzowała się asymetrią prawostronną, z reguły o silnym lub bardzo silnym natężeniu. Oznacza to przewagę powiatów o wartościach cech poniżej średniej arytmetycznej.

współczynników korelacji doprowadziło do uzyskania zbioru cech: $X_2, X_3, X_7, X_8, X_{11}, X_{12}, X_{13}, X_{15}, X_{16}, X_{17}, X_{18}, X_{21}, X_{22}$.

Wykorzystując otrzymane zbiory cech diagnostycznych dokonano liniowego porządkowania powiatów, wykorzystując klasyczny i pozycyjny miernik taksonomiczny. W tabeli 1 zaprezentowano ranking powiatów województwa zachodniopomorskiego pod względem atrakcyjności turystycznej, uzyskany za pomocą klasycznego i pozycyjnego miernika taksonomicznego oraz dwóch różnych metod doboru cech. Rozpatrzono cztery warianty, które otrzymano przy wykorzystaniu:

- wariant I – klasycznego miernika taksonomicznego i parametrycznej metody doboru cech,
- wariant II – klasycznego miernika taksonomicznego i przy doborze cech za pomocą metody odwróconej macierzy,
- wariant III – pozycyjnego miernika taksonomicznego i parametrycznej metody doboru cech,
- wariant IV – pozycyjnego miernika taksonomicznego i przy doborze cech za pomocą metody odwróconej macierzy.

Tabela 1. Porównanie rezultatów klasycznej i pozycyjnej taksonomicznej analizy zróżnicowania atrakcyjności turystycznej powiatów województwa zachodniopomorskiego w 2011 r.

Powiaty	Wariant I		Wariant II		Wariant III		Wariant IV	
	wartość miernika	lokata	wartość miernika	lokata	wartość miernika	lokata	wartość miernika	lokata
Białogardzki	0,1507	16	0,1921	7	0,1846	16	0,1713	11
Drawski	0,2105	10	0,2102	4	0,1655	17	0,0902	18
Kołobrzeski	0,3653	3	0,1754	11	0,4513	3	0,1128	13
Koszaliński	0,1299	18	0,1710	12	0,2021	13	0,2243	8
Sławieński	0,2166	9	0,1795	10	0,2460	9	0,2399	6
Szczecinecki	0,2383	7	0,1891	9	0,2174	11	0,2432	5
Świdwiński	0,1617	13	0,1499	14	0,1600	18	0,2256	7
Wałecki	0,1778	12	0,1242	15	0,2387	10	0,1570	12
M. Koszalin	0,3842	2	0,2663	1	0,5067	2	0,3986	1
Choszczeński	0,2262	8	0,2578	2	0,3776	6	0,3733	2
Gryficki	0,3222	5	0,1636	13	0,3922	5	0,0915	17
Łobeski	0,1471	17	0,1172	16	-0,0353	21	-0,0579	21
Mysłiborski	0,1552	15	0,2363	3	0,3043	8	0,2449	4
Pyrzycki	-0,0475	21	0,1912	8	-0,0066	20	0,1808	10
Stargardzki	0,1570	14	0,1167	17	0,0839	19	-0,0400	20
M. Szczecin	0,1845	11	0,0596	20	0,1878	14	0,1042	16
Goleniowski	0,1245	19	0,1037	18	0,1847	15	0,1097	14
Gryfiński	0,3309	4	0,2005	5	0,4417	4	0,3481	3
Kamiński	0,4031	1	0,1969	6	0,5075	1	0,2147	9
Policki	0,1082	20	-0,0919	21	0,2037	12	-0,0247	19
M. Świnoujście	0,2937	6	0,0600	19	0,3475	7	0,1075	15

Źródło: obliczenia własne na podstawie danych Urzędu Statystycznego w Szczecinie [US 2011].

Jak wynika z tabeli 1 uszeregowania powiatów przy zastosowaniu poszczególnych wariantów nie są jednakowe, a w niektórych przypadkach różnią się dość znacznie. W celu stwierdzenia, czy występują zgodności uporządkowania badanych obiektów przy zastosowaniu omawianych wariantów, wyznaczono współczynniki tau Kendalla (tab. 2)².

Tabela 2. Współczynniki korelacji tau Kendalla obliczone dla rang województw według taksonomicznych mierników rozwoju otrzymanych za pomocą różnych metod doboru cech

Warianty	I	II	III	IV
I	1,0000	0,2857	0,6476	0,2476
II	0,2857	1,0000	0,2571	0,5619
III	0,6476	0,2571	1,0000	0,3714
IV	0,2476	0,5619	0,3714	1,0000

Źródło: obliczenia własne na podstawie tab. 1.

Najwyższe wartości współczynników korelacji uzyskano dla wariantów, w których do konstrukcji taksonomicznych mierników rozwoju wykorzystano zbiory cech uzyskane tymi samymi metodami (wariant I i III oraz II i IV). Wysokie wartości współczynników tau Kendalla świadczą o dość dobrej zgodności liniowego porządkowania powiatów, mimo że występują rozbieżności w pozycjach zajmowanych przez niektóre powiaty, np. powiat policki w porządkowaniu według wariantu I zajmuje 20 miejsce, natomiast według wariantu III – miejsce 12. Dość duże różnice można zauważyć także w przypadku takich powiatów, jak: myśliborski (wariant I – miejsce 15, wariant III – miejsce 8) i drawski (wariant I – miejsce 10, wariant III – miejsce 17). Mimo wyraźnej korelacji porządkowania powiatów według wariantu II i IV i tu zauważyć można znaczne rozbieżności w ich pozycjach. Szczególnie jest to widoczne w przypadku powiatu drawskiego, który w rozpatrywanych wariantach porządkowań zajmuje odpowiednio 4 i 18 pozycję.

Wykorzystując metodę trzech średnich i trzech median zbiorów powiatów podzielono na cztery grupy typologiczne, obejmujące obiekty podobne pod względem atrakcyjności turystycznej (tab. 3).

Mimo że składy poszczególnych klas częściowo się pokrywają (np. m. Koszalin, niezależnie od metody doboru cech oraz konstrukcji miernika taksonomicznego, zawsze znajduje się w grupie pierwszej), to i tak otrzymane grupy różnią się od siebie dość znacznie pod względem przynależności powiatów. Na ogół różnica pomiędzy dwoma różnymi klasyfikacjami tego samego obiektu nie przekracza jednego skupienia, ale trafiają się także znaczniejsze rozbieżności. Na przykład powiat drawski według wariantu drugiego zaliczany jest do skupienia I, natomiast wariant trzeci i czwarty klasyfikują go w grupie IV. Podobnie wygląda sytuacja w przypadku m. Świnoujście – według wariantu pierwszego i trzeciego znajduje

² Współczynniki tau Kendalla przyjmują wartości z przedziału [-1, 1]. Im ich wartość jest bliższa 1, tym większa jest zgodność uporządkowań [Stanisz 2006, s. 313-314].

Tabela 3. Wyniki grupowania powiatów województwa zachodniopomorskiego pod względem atrakcyjności turystycznej w 2011 r.

Grupy	Wariant I	Wariant II	Wariant III	Wariant IV
I	kamieński, m. Koszalin, kołobrzeski, gryfiński, gryficki	m. Koszalin, choszczeński, myśliborski, drawski	kamieński, m. Koszalin, kołobrzeski, gryfiński, gryficki	m. Koszalin, choszczeński, gryfiński, myśliborski, szczecinecki
II	m. Świnoujście, szczecinecki, choszczeński	gryfiński, kamieński, białogardzki, pyrzycki, szczecinecki, sławieński, kołobrzeski, koszaliński, gryficki	choszczeński, m. Świnoujście, myśliborski, sławieński, wałecki, szczecinecki	sławieński, świdwiński, koszaliński, kamieński, pyrzycki, białogardzki
III	sławieński, drawski, m. Szczecin, wałecki, świdwiński, stargardzki, myśliborski, białogardzki, łobeski	świdwiński, wałecki, łobeski, stargardzki, goleniowski	policki, koszaliński, m. Szczecin, goleniowski, białogardzki	wałecki, kołobrzeski, goleniowski, m. Świnoujście, m. Szczecin
IV	koszaliński, goleniowski, policki, pyrzycki	m. Świnoujście, m. Szczecin, policki	drawski, świdwiński, stargardzki, pyrzycki, łobeski	gryficki, drawski, policki, stargardzki, łobeski

Źródło: obliczenia własne na podstawie tab. 1.

się on w grupie II, a według wariantu drugiego – w grupie IV. Aby określić jakość otrzymanych grupowań, zweryfikowano je, wyznaczając wskaźniki homogeniczności, heterogeniczności oraz poprawności skupień (tab. 4).

Tabela 4. Mierniki oceny poprawności grupowania obiektów

Wskaźniki	Wariant I	Wariant II	Wariant III	Wariant IV
Homogeniczność skupień	0,4608	0,2045	0,3985	0,2306
Heterogeniczność skupień	0,7421	0,7352	0,8619	0,2825
Poprawność skupień	0,6209	0,2781	0,4624	0,8160

Źródło: obliczenia własne na podstawie tab. 1.

Analizując wyniki dotyczące jakości grupowań, można stwierdzić, że pod względem homogeniczności i poprawności grupowania najlepsza okazała się klasyfikacja wykorzystująca klasyczny miernik syntetyczny, gdy wybór zmiennych dokonany został za pomocą metody odwróconej macierzy (wariant II). Natomiast w zakresie heterogeniczności skupień lepsze rezultaty otrzymano dla grupowania powiatów z wykorzystaniem miernika pozycyjnego i parametrycznej metody doboru cech (wariant III).

W klasyfikacji przeprowadzonej według wariantu drugiego w najlepszej pierwszej grupie znalazły się następujące powiaty: m. Koszalin, choszczeński, myśliborski, drawski. Charakteryzują się one korzystnymi wielkościami średnich w porównaniu ze średnimi ogólnymi, odnoszącymi się do następujących cech: lesistość,

łóżka w szpitalach na 10 tys. ludności, miejsca na widowni w kinach na 1000 ludności, muzealia na 1000 ludności. W grupie tej jest najmniejsza emisja zanieczyszczeń gazowych powietrza w t na 1 km² oraz odpady wytworzone w ciągu roku ogółem w tys. t na 1 km².

Dobra sytuacja w drugiej grupie powiatów występuje w przypadku takich średnich wartości cech, jak: pomniki przyrody na 1 km², drogi publiczne powiatowe o twardej nawierzchni w km na 100 km², targowiska stałe na 1000 ludności, stałe punkty sprzedaży drobnodetalicznej na 1000 ludności. Do niekorzystnych wartości należą średnie dotyczące liczby ludności na 1 aptekę, miejsc na widowni w kinach na 1000 ludności.

Słaba atrakcyjność turystyczna powiatów wchodzących w skład grupy trzeciej wynika przede wszystkim ze słabego zagospodarowania turystycznego (mała liczba: targowisk stałych na 1000 ludności, stałych punktów sprzedaży drobnodetalicznej na 1000 ludności, muzealiów na 1000 ludności). Pozytywny wpływ na badane zjawisko mają: dość duża lesistość oraz mała ilość odpadów wytworzonych w ciągu roku ogółem w tys. ton na 1 km².

Grupa czwarta to zbiór powiatów, dla których większość cech przyjmuje niekorzystne wartości w porównaniu ze średnimi dla całego województwa. Negatywny wpływ na atrakcyjność turystyczną w tej klasie ma niska średnia wartość odnosząca się do: pomników przyrody, lesistości, targowisk stałych na 1000 ludności. Ponadto w klasie tej zaobserwowano najwyższą ilość odpadów wytworzonych na 1 km² oraz emisji zanieczyszczeń gazowych powietrza w tonach na 1 km².

4. Podsumowanie

W artykule przedstawiono zastosowanie klasycznego i pozycyjnego podejścia w problematyce porządkowania liniowego, przy jednoczesnym wykorzystaniu dwóch różnych metod doboru cech diagnostycznych: parametrycznej i odwróconej macierzy. Zbadano także wpływ wyników otrzymanych w poszczególnych wariantach na jakość klasyfikacji. Przedstawione podejście zilustrowano przykładem dotyczącym klasyfikacji powiatów województwa zachodniopomorskiego pod względem atrakcyjności turystycznej w 2011 roku. Otrzymane porządki różnią się między sobą, niewiele powiatów zajmuje zbliżone pozycje w klasyfikacji, a tylko jeden obiekt (m. Koszalin) w przypadku każdego wariantu należy do tej samej najlepszej pierwszej grupy.

Wskaźniki określające jakość grupowań wskazują na to, że grupowanie dokonane na podstawie wartości miernika klasycznego, gdy wybór zmiennych dokonany został za pomocą metody odwróconej macierzy (wariant II), cechuje się lepszą homogenicznością i poprawnością skupień. Odbywa się to jednak kosztem pogorszenia ich heterogeniczności.

Na podstawie przeprowadzonego badania można stwierdzić, że wyniki uzyskane za pomocą różnych metod doboru cech i różnych mierników taksonomicznych (klasyczny, pozycyjny) mają wpływ na jakość grupowań.

Literatura

- Grabiński T. (1992), *Metody taksonometrii*, Akademia Ekonomiczna w Krakowie, Kraków.
- Młodak A. (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa.
- Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa.
- Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, Oficyna Wydawnicza SGGW, Warszawa.
- Rapacz A. (red.) (2012), *Wyzwania współczesnej polityki turystycznej. Problemy funkcjonowania rynku turystycznego*, Prace Naukowe UE we Wrocławiu nr 258, Wydawnictwo UE we Wrocławiu, Wrocław.
- Stanisz A. (2006), *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*, t. 1. *Statystyki podstawowe*, Wydawnictwo StatSoft Polska, Kraków.
- US (2011), *Województwo zachodniopomorskie, podregiony, powiaty, gminy*, Urząd Statystyczny w Szczecinie, Szczecin.
- Wysocki F. (2010), *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*, Wyd. Uniwersytetu Przyrodniczego w Poznaniu, Poznań.

THE COMPARISON OF THE QUALITY OF GROUPINGS OF POWIATS OF WEST POMERANIAN VOIVODESHIP IN TERMS OF TOURISM ATTRACTIVENESS

Summary: The main goal of this article is to compare the quality of groupings of poviats of West Pomeranian Voivodeship in terms of their tourism attractiveness, using: different methods of selection of diagnostic variables (parametric method, the method of inverse matrix) and classical and positional taxonomic indicators of development. In the study there were used diagnostic variables characterizing the tourist valorization, the level of tourism development, transport accessibility and the level of environmental pollution. The quality of groupings procedures was checked by calculating indicators of homogeneity, heterogeneity and clustering accuracy.

Keywords: method of diagnostic variables selection, Hellwig's development indicator, Weber's median, quality of groupings, tourism attractiveness.