

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

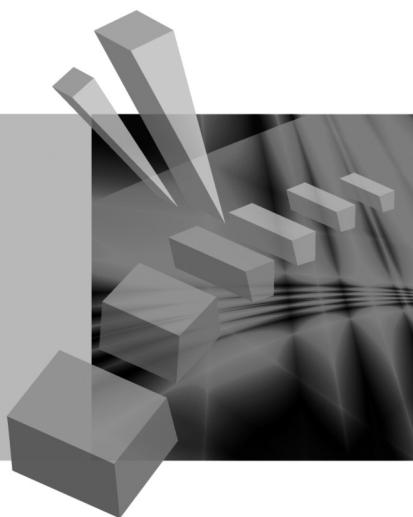
of Wrocław University of Economics

242

Taksonomia 19.

Klasyfikacja i analiza danych

– teoria i zastosowania



Redaktorzy naukowi
Krzysztof Jajuga
Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2012

Recenzenci: Eugeniusz Gatnar, Elżbieta Gołata, Tadeusz Kufel, Józef Pocięcha,
Miroslaw Szreder, Feliks Wysocki

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Tytuł sfinansowano ze środków Sekcji Klasyfikacji i Analizy Danych PTS
i Uniwersytetu Ekonomicznego we Wrocławiu

Publikacja jest dostępna na stronie www.ibuk.pl

Streszczenia opublikowanych artykułów są dostępne w międzynarodowej bazie danych
The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl>
oraz w 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](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

Kopowanie i powielanie w jakiegokolwiek formie
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2012

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM
Nakład: 320 egz.

Spis treści

Wstęp	13
Stanisława Bartosiewicz , Jeszcze raz o skutkach subiektywizmu w analizie wielowymiarowej	17
Andrzej Sokolowski , Q uniwersalna miara odległości	22
Eugeniusz Gatnar , Jakość danych w systemach statystycznych banków centralnych (na przykładzie NBP)	31
Marek Walesiak , Pomiar odległości obiektów opisanych zmiennymi mierzonymi na skali porządkowej – strategię postępowania.....	39
Krzysztof Jajuga, Marek Walesiak , XXV lat konferencji taksonomicznych – fakty i refleksje	47
Józef Pocięcha, Barbara Pawelek , Model SEM w analizie zagrożenia bankructwem przedsiębiorstw w świetle koniunktury gospodarczej – problemy teoretyczne i praktyczne	50
Paweł Lula , Uczące się systemy pozyskiwania informacji z dokumentów tekstowych	58
Ewa Roszkowska , Zastosowanie metody TOPSIS do wspomagania procesu negocjacji.....	68
Andrzej Młodak , Sąsiedztwo obszarów przestrzennych w ujęciu fizycznym oraz społeczno-ekonomicznym – podejście taksonomiczne	76
Andrzej Bąk , Modele kategorii nieuporządkowanych w badaniach preferencji	86
Jacek Kowalewski , Zintegrowany model optymalizacji badań statystycznych.....	96
Jan Paradysz, Karolina Paradysz , Obszary bezrobocia w Polsce – problem benchmarkowy.....	106
Tomasz Szubert , W co grać, aby jak najmniej przegrać? Próba klasyfikacji systemów gry w zakładach bukmacherskich.....	116
Izabela Szamrej-Baran , Klasyfikacja krajów UE ze względu na ubóstwo energetyczne	126
Sylwia Filas-Przybył, Tomasz Klimanek, Jacek Kowalewski , Analiza dojazdów do pracy za pomocą modelu grawitacji.....	135
Marta Dziechciarz-Duda, Anna Król, Klaudia Przybysz , Minimum egzystencji a czynniki warunkujące skłonność do korzystania z pomocy społecznej. Klasyfikacja gospodarstw domowych	144
Hanna Dudek , Subiektywne skale ekwiwalentności – analiza na podstawie danych o satysfakcji z osiągniętych dochodów	153

Joanicjusz Nazarko, Ewa Chodakowska, Marta Jaročka, Segmentacja szkół wyższych metodą analizy skupień <i>versus</i> konkurencja technologiczna ustalona metodą DEA – studium komparatywne.....	163
Ewa Chodakowska, Wybrane metody klasyfikacji w konstrukcji ratingu szkół.....	173
Bartosz Soliński, Sektor energetyki odnawialnej w krajach Unii Europejskiej – klasyfikacja w świetle strategii zarządzania zmianą.....	182
Krzysztof Szwarz, Klasyfikacja powiatów województwa wielkopolskiego ze względu na sytuację demograficzną.....	192
Elżbieta Gołata, Grażyna Dehnel, Rejestry administracyjne w analizie przedsiębiorczości.....	202
Katarzyna Chudy, Marek Sobolewski, Kinga Stępień, Wykorzystanie metod taksonomicznych w prognozowaniu wskaźników rentowności banków giełdowych w Polsce.....	212
Katarzyna Dębowska, Modelowanie upadłości przedsiębiorstw przy wykorzystaniu metod dyskryminacji i regresji.....	222
Alina Bojan, Wykorzystanie metod wielowymiarowej analizy danych do identyfikacji zmiennych wpływających na atrakcyjność wybranych inwestycji.....	231
Justyna Brzezińska, Analiza logarytmiczno-liniowa w badaniu przyczyn umieralności w krajach UE.....	240
Aneta Rybicka, Bartłomiej Jefmański, Marcin Pelka, Analiza klas ukrytych w badaniach satysfakcji studentów.....	247
Bartłomiej Jefmański, Pomiar opinii respondentów z wykorzystaniem elementów teorii zbiorów rozmytych i środowiska R.....	256
Julita Stańczuk, Porównanie rezultatów wielostanowej klasyfikacji obiektów ekonomicznych z wykorzystaniem analizy dyskryminacyjnej oraz sieci neuronowych.....	265
Jerzy Krawczuk, Skuteczność metod klasyfikacji w prognozowaniu kierunku zmian indeksu giełdowego S&P500.....	275
Anna Czapkiewicz, Beata Basiura, Symulacyjne badanie wpływu zaburzeń na grupowanie szeregów czasowych na podstawie modelu Copula-GARCH.....	283
Radosław Pietrzyk, Ocena efektywności inwestycji funduszy inwestycyjnych z tytułu doboru papierów wartościowych i umiejętności wykorzystania trendów rynkowych.....	291
Aleksandra Witkowska, Marek Witkowski, Zastosowanie metody Panzara-Rosse’a do pomiaru poziomu konkurencji w sektorze banków spółdzielczych.....	306
Marcin Pelka, Podejście wielomodelowe z wykorzystaniem metody <i>boosting</i> w analizie danych symbolicznych.....	315
Justyna Wilk, Analiza porównawcza oprogramowania komputerowego w klasyfikacji danych symbolicznych.....	323

Tomasz Bartłomowicz, Justyna Wilk , Zastosowanie metod analizy danych symbolicznych w przeszukiwaniu dziedzinowych baz danych.....	333
Kamila Migdał-Najman , Propozycja hybrydowej metody grupowania opartej na sieciach samouczących	342
Dorota Rozmus , Porównanie dokładności taksonomii spektralnej oraz zagregowanych algorytmów taksonomicznych opartych na idei metody <i>bagging</i>	352
Krzysztof Najman , Grupowanie dynamiczne z wykorzystaniem samouczących się sieci GNG	361
Małgorzata Misztal , Wpływ wybranych metod uzupełniania brakujących danych na wyniki klasyfikacji obiektów z wykorzystaniem drzew klasyfikacyjnych w przypadku zbiorów danych o niewielkiej liczebności – ocena symulacyjna	370
Mariusz Kubus , Zastosowanie wstępnego uwarunkowania zmiennej objaśnianej do selekcji zmiennych.....	380
Barbara Batóg, Jacek Batóg , Wykorzystanie analizy dyskryminacyjnej do identyfikacji czynników determinujących stopę zwrotu z inwestycji na rynku kapitałowym	387
Katarzyna Wójcik, Janusz Tuchowski , Analiza porównawcza miar podobieństwa tekstów opartych na macierzy częstości i tekstów opartych na wiedzy dziedzinowej	396
Iwona Staniec , Analiza czynnikowa w identyfikacji obszarów determinujących doskonalenie systemów zarządzania w polskich organizacjach	406
Marek Lubicz, Maciej Zięba, Adam Rzechonek, Konrad Pawelczyk, Jerzy Kołodziej, Jerzy Błaszczyk , Analiza porównawcza wybranych technik eksploracji danych do klasyfikacji danych medycznych z brakującymi obserwacjami	416
Iwona Foryś , Wykorzystanie analizy log-liniowej do wyboru czynników determinujących atrakcyjność cenową mieszkań w obrocie wtórnym na przykładzie lokalnego rynku mieszkaniowego.....	426
Ewa Genge , Analiza skupień oparta na mieszankach uciętych rozkładów normalnych.....	436
Jerzy Korzeniewski , Ocena efektywności metody uśredniania zmiennych i metody Ichino selekcji zmiennych w analizie skupień	444
Andrzej Dudek , SMS – propozycja nowego algorytmu analizy skupień	451
Artur Mikulec , Metody oceny wyniku grupowania w analizie skupień.....	460
Małgorzata Machowska-Szewczyk , Algorytm klasyfikacji rozmytej dla obiektów opisanych za pomocą zmiennych symbolicznych oraz rozmytych	469
Artur Zaborski , Analiza PROFIT i jej wykorzystanie w badaniu preferencji	479
Karolina Bartos , Analiza skupień wybranych państw ze względu na strukturę wydatków konsumpcyjnych obywateli – zastosowanie sieci Kohonena	488

Barbara Batóg, Magdalena Mojsiewicz, Katarzyna Wawrzyniak , Klasyfikacja gospodarstw domowych ze względu na bodźce do zawierania umowy o ubezpieczenie z wykorzystaniem modeli zmiennych jakościowych .	496
Izabela Kurzawa , Zastosowanie modelu LA/AIDS do badania elastyczności cenowych popytu konsumpcyjnego w gospodarstwach domowych w relacji miasto–wieś	505
Aleksandra Łuczak, Feliks Wysocki , Metody porządkowania liniowego obiektów opisanych za pomocą cech metrycznych i porządkowych	513
Agnieszka Sompolska-Rzechuła , Porównanie klasycznej i pozycyjnej taksonomicznej analizy zróżnicowania jakości życia w województwie zachodniopomorskim	523
Joanna Banaś, Małgorzata Machowska-Szewczyk , Ocena intensywności wykorzystania skrzynek poczty elektronicznej za pomocą uporządkowanego modelu probitowego	532
Iwona Bąk , Segmentacja gospodarstw domowych emerytów i rencistów pod względem wydatków na rekreację i kulturę	541
Aneta Becker , Zastosowanie metody ANP do porządkowania województw Polski pod względem dynamiki wykorzystania ICT w latach 2008-2010	552
Katarzyna Dębowska , Klasyfikacja sektorów ze względu na ich kondycję finansową przy użyciu metod wielowymiarowej analizy statystycznej	562
Anna Domagała , Propozycja metody doboru zmiennych do modeli DEA (procedura kombinowanego doboru w przód).....	571
Henryk Gierszal, Karina Pawlina, Maria Urbańska , Analiza statystyczna w badaniach zapotrzebowania na usługi teleinformatyczne sieci łączności ruchomej	580
Hanna Gruchociak , Konstrukcja estymatora regresyjnego dla danych o strukturze dwupoziomowej.....	590
Tomasz Klimanek, Marcin Szymkowiak , Zastosowanie estymacji pośredniej uwzględniającej korelację przestrzenną w opisie niektórych charakterystyk rynku pracy	601
Jarosław Lira , Prognozowanie opłacalności produkcji żywca wieprzowego w Polsce	610
Christian Lis , Wykorzystanie metody klasyfikacji w ocenie konkurencyjności portów południowego Bałtyku	619
Beata Bieszk-Stolorz, Iwona Markowicz , Wykorzystanie wielomianowego modelu logitowego do oceny szansy podjęcia pracy przez bezrobotnych .	628
Lucyna Przezbórska-Skobiej, Jarosław Lira , Przestrzeń agroturystyczna Polski i ocena jej atrakcyjności.....	637
Paweł Ulman , Model rozkładu wydatków a funkcje popytu.....	646
Maria Urbańska, Tadeusz Mizera, Henryk Gierszal , Zastosowanie metod analizy statystycznej w badaniach mięczaków	655

Summaries

Stanisława Bartosiewicz , The effects of subjectivism in multivariate analysis revisited.....	21
Andrzej Sokółowski , Q universal distance measure	30
Eugeniusz Gatnar , Data quality in central banks' statistical systems (NBP example)	38
Marek Walesiak , Distance measures for ordinal data – strategies of proceedings.....	46
Krzysztof Jajuga, Marek Walesiak , XXV years of taxonomic conferences – some facts and remarks.....	49
Józef Pocięcha, Barbara Pawelek , General SEM model in researching corporate bankruptcy and business cycles – theoretical and practical problems.....	57
Paweł Lula , Learning-based systems of information extraction from textual resources	67
Ewa Roszkowska , The application of the TOPSIS method to support the negotiation process	75
Andrzej Młodak , Neighborhood of spatial areas in the physical and socio-economic context – a taxonomic approach.....	85
Andrzej Bąk , Models for unordered categories in preference analysis.....	95
Kowalewski Jacek , An integrated model of optimizing statistical surveys	105
Jan Paradysz, Karolina Paradysz , Areas of unemployment in Poland – benchmark problem	115
Tomasz Szubert , How to play to lose the least? Classification of systems in sports bets	125
Izabela Szamrej-Baran , Classification of EU member states in view of fuel poverty	134
Sylvia Filas-Przybył, Tomasz Klimanek, Jacek Kowalewski , An attempt to use the gravity model in the analysis of commuters.....	143
Marta Dziechciarz-Duda, Anna Król, Klaudia Przybysz , Subsistence minimum versus factors influencing tendency to benefit from social care. Classification of households	152
Hanna Dudek , Subjective equivalence scales – analysis based on data about satisfaction with incomes.....	162
Joanicjusz Nazarko, Ewa Chodakowska, Marta Jarocka , Segmentation of universities using cluster analysis versus technological competitors determined by the DEA method – a comparative study	172
Ewa Chodakowska , Selected methods of classification in schools' rating.....	181
Bartosz Soliński , Renewable energy sector in the European Union – classification in the light of change management strategy	191
Krzysztof Szwarz , Classification of Wielkopolska voivodeship due to the demographic situation	201

Elżbieta Gołata, Grażyna Dehnel , Administrative registers in business analysis.....	211
Katarzyna Chudy, Marek Sobolewski, Kinga Stępień , Application of taxonomic methods in forecasting the profitability ratios of listed banks in Poland.....	221
Katarzyna Dębowska , Modeling bankruptcy of firms by using discrimination and regression methods.....	230
Alina Bojan , Identification of variables which influence attractiveness of given investments with the usage of multivariate analysis.....	239
Justyna Brzezińska , Log-linear analysis in the study of mortality in EU.....	246
Aneta Rybicka, Bartłomiej Jefmański, Marcin Pelka , Latent class analysis in student satisfaction surveys.....	254
Bartłomiej Jefmański , The respondent's opinions measurement in the R program with an application of fuzzy sets theory.....	264
Julita Stańczuk , A comparison of the results of multistate classification of economic objects using discriminant analysis and artificial neural networks.....	274
Jerzy Krawczuk , Effectiveness of classification methods in S&P500 stock index direction changes forecasting.....	282
Anna Czapkiewicz, Beata Basiura , The simulation study of the utility of the Copula-GARCH models for clustering financial time series.....	290
Radosław Pietrzyk , Timing and selectivity in mutual funds performance measurement.....	305
Aleksandra Witkowska, Marek Witkowski , Use of the Panzar-Rosse method to assess of the competition level in the cooperative banks sector.....	314
Marcin Pelka , Ensemble learning with the application of <i>boosting</i> in symbolic data analysis.....	322
Justyna Wilk , Comparative study of symbolic data classification software.....	332
Tomasz Bartłomowicz, Justyna Wilk , Application of symbolic data analysis methods for domain database searching.....	341
Kamila Migdał-Najman , A proposal of hybrid clustering method based on self-learning networks.....	351
Dorota Rozmus , Comparison of accuracy of spectral clustering and cluster ensembles stability based on bagging idea.....	360
Krzysztof Najman , A dynamic grouping based on self-learning GNG networks.....	369
Małgorzata Misztal , Influence of data imputation methods on the results of object classification using classification trees in the case of small data sets – simulation assessment.....	379
Mariusz Kubus , The application of pre-conditioning of explanatory variable for feature selection.....	386
Barbara Batóg, Jacek Batóg , Application of discriminant analysis to the identification of factors determining the rate of return on the capital market.....	395

Katarzyna Wójcik, Janusz Tuchowski , Comparative analysis of text documents similarity measures based on frequency matrix and based on domain knowledge.....	405
Iwona Staniec , Factor analysis in the identification of areas that determine the improvement of management systems in Polish organizations.....	415
Marek Lubicz, Maciej Zięba, Adam Rzechonek, Konrad Pawełczyk, Jerzy Kołodziej, Jerzy Błaszczyk , Comparative analysis of selected data mining approaches to the classification of medical data with missing values (covariates).....	425
Iwona Foryś , The log-linear analysis using to select the factors determining the attractiveness of the price of flats on the secondary market on the example of local housing market.....	435
Ewa Genge , Trimming approach to the mixtures of normal distributions.....	443
Jerzy Korzeniewski , Efficiency assessment of Ichino method and mean value method of selecting variables in cluster analysis.....	450
Andrzej Dudek , SMS – proposal of new clustering algorithm.....	459
Artur Mikulec , Evaluation methods for the grouping result in cluster analysis.....	468
Małgorzata Machowska-Szewczyk , Fuzzy clustering algorithm for objects described by symbolic or fuzzy variables.....	478
Artur Zaborski , PROFIT analysis and its using in the research of preferences.....	487
Karolina Bartos , Cluster analysis of selected countries due to the structure of their citizens' consumer expenditures – the use of Kohonen networks.....	495
Barbara Batóg, Magdalena Mojsiewicz, Katarzyna Wawrzyniak , Classification of households according to the impulses of concluding the insurance contract by means of qualitative variable models.....	504
Izabela Kurzawa , The application of LA/AIDS model to examine price elasticities of demand of households in the urban-rural relationship.....	512
Aleksandra Luczak, Feliks Wysocki , Linear ordering methods of objects described by a set of metric and ordinal characteristics.....	522
Agnieszka Sompolska-Rzechuła , The comparison of the classical and positional taxonomic analysis of the quality of life differentiation in Zachodniopomorskie voivodeship.....	531
Joanna Banaś, Małgorzata Machowska-Szewczyk , Evaluation of intensity of mailboxes using with the ordered probit model.....	540
Iwona Bąk , Segmentation of pensioners and annuitants households in terms of expenditures on recreation and culture.....	551
Aneta Becker , Application of ANP method to organize Polish voivodships in terms of dynamics of the use of ICT in 2008-2010.....	561
Katarzyna Dębowska , The classification of sectors' financial situation using the methods of multivariate statistical analysis.....	570

Anna Domagała , Proposal of a new method for variable selection in DEA models (combined forward stepwise selection method).....	579
Henryk Gierszal, Karina Pawlina, Maria Urbańska , Statistical analysis in demand research of ICT services in mobile networks.....	589
Hanna Gruchociak , Construction of regression estimator for two-level data	600
Tomasz Klimanek, Marcin Szymkowiak , Application of spatial models in indirect estimation of some labor market characteristics	609
Jarosław Lira , Forecasting of hog livestock production profitability in Poland	618
Christian Lis , The utilization of taxonomic methods in the appraisal of competitiveness of south Baltic ports	627
Beata Bieszk-Stolorz, Iwona Markowicz , The application of the multinomial logit model in evaluating employment odds for the unemployed job seekers	636
Lucyna Przezbórska-Skobiej, Jarosław Lira , Agritourism space of Poland and its valuation.....	645
Paweł Ulman , Model of expenses distribution and demand functions.....	654
Maria Urbańska, Tadeusz Mizera, Henryk Gierszal , Methods of statistical analysis in research of molluscs	663

Agnieszka Sompolska-Rzechuła

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

PORÓWNANIE KLASYCZNEJ I POZYCYJNEJ TAKSONOMICZNEJ ANALIZY ZRÓŻNICOWANIA JAKOŚCI ŻYCIA W WOJEWÓDZTWIE ZACHODNIOPOMORSKIM

Streszczenie: Celem artykułu było porównanie efektywności dwóch podejść w konstrukcji miernika rozwoju wykorzystujących klasyczny miernik Hellwiga oraz pozycyjny wektor medianowy Webera w badaniu obiektywnej jakości życia w województwie zachodniopomorskim. W analizie wykorzystano wskaźniki z następujących dziedzin: środowisko naturalne, demografia, rynek pracy, infrastruktura komunalna i mieszkania, edukacja, kultura i turystyka, ochrona zdrowia, transport drogowy, dochody i wydatki budżetów jednostek samorządu terytorialnego. Badanie dotyczyło 2008 r. Podejście klasyczne dało lepsze wyniki pod względem homogeniczności, heterogeniczności i poprawności grupowania.

Słowa kluczowe: mediana Webera, jakość życia, porządkowanie liniowe.

1. Wstęp

Dotychczas w różnych pracach naukowych powstało bardzo wiele definicji jakości. Najprostszą, a zarazem najbardziej uniwersalną, jest ta, z której wynika, że „Jakość jest to stopień spełnienia stawianych wymagań” [Kolman 2009, s. 142]. W każdej dziedzinie działalności ludzkiej formułowane są różne wymagania, zatem z jakością człowiek ma do czynienia wielokrotnie i na co dzień.

Jakość życia jest pojęciem złożonym i podobnie jak wiele innych pojęć stosowanych w naukach społecznych niemającym jednej, powszechnie akceptowanej definicji. Ogólnie można powiedzieć o trzech głównych podejściach do statystycznego pomiaru jakości życia. Po pierwsze, można przyjąć, że jakość życia jest to poziom zadowolenia, jaki uzyskuje człowiek z różnych sfer życia. Według tej koncepcji, pomiaru jakości życia dokonuje się na podstawie deklarowanych bezpośrednio przez jednostkę (respondenta) ocen i opinii, dotyczących m.in.: stopnia zaspokojenia różnego rodzaju potrzeb, poziomu satysfakcji z aktywności w rozmaitych dziedzinach życia czy poczucia szczęścia. Po drugie, można wziąć pod uwagę jedynie tzw. obiektywne wskaźniki, charakteryzujące zarówno materialne, jak i niematerialne aspekty warunków życia. Każde z tych podejść ma swoje zalety i ograniczenia i każde

z osobna nie pozwala na dokonanie kompleksowej (całościowej) oceny jakości życia. Optymalnym rozwiązaniem wydaje się trzecie podejście, w którym wykorzystuje się obie kategorie miar. Integracja danych o charakterze obiektywnym i subiektywnym w znacznym stopniu zwiększa możliwości analityczne, a tym samym praktyczne wykorzystanie wyników tych analiz zarówno do potrzeb kreowania polityki społeczno-ekonomicznej, jak i do monitorowania zmian dobrobytu społeczno-ekonomicznego [Szukielojć-Bieńkuńska, Walczak 2011]. W miarę możliwości powinno dążyć się do uwzględnienia pomiaru w ujęciu regionalnym.

W związku z tak dużym zainteresowaniem tą kategorią istnieje wiele określeń jakości życia. Istota jakości życia i pojęć z niej wynikających są związane z kryteriami definiowana tej kategorii [Borys 2008]. Można wymienić następujące kryteria klasyfikacji jakości życia: wartościowania (oceny), zakresu i liczby aspektów, obiektywności pomiaru, liczby obiektów, bezpośredniości powiązania badań z rzeczywistą jakością życia, zrównoważenia aspektów jakości życia, ujawniania w jakości życia systemu wartości.

W pracy podjęto próbę analizy jakości życia w odniesieniu do kryterium obiektywności.

Obiektywną jakość życia określa się często jako warunki życia lub poziom życia. Subiektywna jakość życia była dotychczas domeną przedstawicieli nauk socjologicznych, którzy wykorzystywali do określenia istoty jakości życia takie terminy, jak zadowolenie, satysfakcja, optymizm. Przykładem takiej definicji jest określenie jakości życia jako stopnia satysfakcji (zadowolenia) człowieka z całej swojej egzystencji [Ostasiewicz 2004].

Celem artykułu jest porównanie efektywności dwóch podejść w konstrukcji miernika rozwoju wykorzystujących klasyczny miernik Hellwiga oraz pozycyjny wektor medianowy Webera w badaniu jakości życia w województwie zachodniopomorskim.

2. Metoda i materiał badawczy

W ocenie jakości życia w ujęciu obiektywnym bardzo często znajdują zastosowanie zmienne syntetyczne. Wyznaczona zmienna pozwala na uporządkowanie i porównanie analizowanych obiektów, w tym przypadku powiatów województwa zachodniopomorskiego, można ją także wykorzystać do wyznaczenia grup powiatów charakteryzujących się podobną obiektywną jakością życia. Istotną sprawą w procedurze konstrukcji zmiennej syntetycznej jest rozróżnienie zmiennych syntetycznych bezwzorcowych i wzorcowych. W pracy do budowy taksonomicznego miernika rozwoju wykorzystano metodę wzorcową, w podejściu zarówno klasycznym, jak i pozycyjnym, wykorzystującym medianę Webera¹. W ujęciu klasycznym budowa taksonomicznego miernika została przedstawiona np. w pracy [Nowak 1991, s. 87-91].

¹ Mediana Webera stanowi wielowymiarowe uogólnienie klasycznego pojęcia mediany. Chodzi tu o wektor, który minimalizuje sumę euklidesowych odległości od danych punktów reprezentujących

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 bezwzględnego z zastosowaniem mediany Webera [Młodak 2006, s. 136-137]:

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

gdzie $\theta_0 = (\theta_{01}, \theta_{02}, \dots, \theta_{0m})$ jest medianą Webera, $\text{m}\ddot{\text{a}}\text{d}(X_j)$ to medianowe odchylenie bezwzględne, w którym bada się dystanse cech do wektora Webera, czyli: $\text{m}\ddot{\text{a}}\text{d}(X_j) = \text{med}_{i=1,2,\dots,n} |x_{ij} - \theta_{0j}|$ ($j = 1, 2, \dots, m$).

Miernik agregatowy wyznacza się według wzoru:

$$\mu_i = 1 - \frac{d_{i0}}{d_0},$$

gdzie $d_0 = \text{med}(\mathbf{d}) + 2,5 \cdot \text{mad}(\mathbf{d})$, przy czym $\mathbf{d} = (d_1, d_2, \dots, d_n)$ – wektor odległości wyznaczany według wzoru: $d_i = \text{med}_{j=1,2,\dots,m} |z_{ij} - \varphi_j|$ $i = 1, 2, \dots, n$, $\varphi_j = \max_{i=1,2,\dots,n} z_{ij}$ – współrzędne wektora wzorca rozwoju, którymi są maksymalne wartości znormalizowanych cech.

Uporządkowania obiektów zarówno za pomocą miernika klasycznego, jak i z wykorzystaniem mediany Webera stanowią podstawę do ich typologii. Grupy typologiczne oparte na podejściu klasycznym wykorzystują średnią arytmetyczną (\bar{z}) oraz odchylenie standardowe (s_z) mierników rozwoju, natomiast w ujęciu pozycyjnym zaproponowanym w pracy [Lira, Wagner, Wysoki 2002, s. 93] klasy typologiczne są następujące:

- grupa I: $\mu_i \geq \text{med}(\mu) + 2,5 \cdot \text{mad}(\mu)$,
- grupa II: $\text{med}(\mu) \leq \mu_i < \text{med}(\mu) + 2,5 \cdot \text{mad}(\mu)$,
- grupa III: $\text{med}(\mu) - 2,5 \cdot \text{mad}(\mu) \leq \mu_i < \text{med}(\mu)$,
- grupa IV: $\mu_i < \text{med}(\mu) - 2,5 \cdot \text{mad}(\mu)$.

Wykorzystanie metody opartej na medianie Webera pozwala zniwelować zakłócający wpływ obserwacji odstających, daje możliwość ustalenia udziału poszczególnych odchyleń w odpowiedniej wartości agregatowej, a przy tym w całym procesie badawczym traktuje zbiór cech diagnostycznych jako jedną całość. W ocenie homogeniczności otrzymanych grup wykorzystano miernik o następującej postaci [Młodak 2006, s. 138-141]:

rozpatrywane obiekty, a więc znajduje się niejako „pośrodku” nich, ale jest jednocześnie uodporniony na występowanie obserwacji odstających.

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

gdzie: $hm_6^*(P_k) = \text{med}_{i: O_i \in P_k} \delta(O_i, \Gamma_{\theta_k})$ – 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,

k – liczba klas, $k = 1, 2, \dots, p$,

p – liczba skupień otrzymanych na danym poziomie grupowania.

Natomiast w ocenie heterogeniczności zastosowano miernik:

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

gdzie $ht_6^*(P_k) = \text{med}_{\substack{i=1, \dots, p \\ i \neq k}} \delta(\Gamma_{\theta_i}, \Gamma_{\theta_k})$ – 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 oceny jakości analizy taksonomicznej stosuje się mierniki homogeniczności oraz heterogeniczności skupień, wykorzystując koncepcję środka ciężkości grupy i odległości od niego. W badaniu wykorzystano podejście, w którym środek ciężkości danej grupy zastąpiony został medianą Webera jej elementów. Zatem mediana Webera ma zastosowanie nie tylko jako narzędzie służące do konstrukcji taksonomicznego miernika rozwoju, ale także w weryfikacji efektywności różnych metod badania różnicowania obiektów wielocechowych.

Źródłem informacji o jakości życia w ujęciu obiektywnym mieszkańców powiatów województwa zachodniopomorskiego była publikacja Urzędu Statystycznego w Szczecinie *Warunki życia w województwie zachodniopomorskim w latach 2006-2008*. Badanie dotyczyło roku 2008. Do pomiaru jakości życia w ujęciu obiektywnym zastosowano wstępny zbiór cech, który obejmował 44 cechy, pochodzące z takich dziedzin życia, jak: środowisko naturalne, uwarunkowania demograficzne, rynek pracy, infrastruktura komunalna i mieszkania, edukacja, kultura i turystyka, ochrona zdrowia, transport drogowy, dochody i wydatki budżetów jednostek samorządu terytorialnego. Po zebraniu niezbędnych danych dokonano ich weryfikacji. Usunięto te cechy, dla których współczynnik zmienności posiadał wartość mniejszą niż 10%, były to cechy: X_{10} – małżeństwa zawarte na 1000 ludności, X_{18} – bezrobot-

ne kobiety w liczbie bezrobotnych ogółem w %, X_{23} – przeciętna liczba osób na 1 mieszkanie, X_{26} – liczba szkół podstawowych z dostępem do Internetu w ogólnej liczbie szkół podstawowych w %, X_{27} – liczba gimnazjów z dostępem do Internetu w ogólnej liczbie gimnazjów w %, X_{28} – liczba liceów ogólnokształcących z dostępem do Internetu w ogólnej liczbie liceów ogólnokształcących w %, X_{39} – samochody osobowe zarejestrowane na 1000 ludności. W kolejnym etapie oceniono wzajemne skorelowanie pozostałych cech. W tym celu wykorzystano metodę odwróconej macierzy współczynników korelacji. W analizowanym przypadku wystąpiła dosyć rzadka sytuacja, w której nie można odwrócić macierzy współczynników korelacji, ponieważ wyznacznik tejże macierzy wynosił zero. W tej sytuacji usuwano te cechy, które charakteryzowały się największą liczbą współczynników korelacji z innymi cechami przekraczającymi wartość 0,60. Postępowanie takie przeprowadzano aż do momentu, w którym macierz współczynników korelacji można było odwrócić i wówczas wyeliminowano te cechy, które odpowiadały nadmiernym elementom diagonalnym tejże macierzy odwrotnej [Słaby 2006, s. 118-131]. W efekcie tych czynności określono zbiór cech diagnostycznych, na który składają się następujące cechy: X_3 – stopień redukcji wytworzonych zanieczyszczeń w %, X_5 – powierzchnia o szczególnych walorach przyrodniczych prawnie chroniona w powierzchni ogólnej w %, X_{12} – zgony niemowląt na 1000 ludności, X_{15} – zgony z powodu chorób układu oddechowego na 1000 ludności, X_{29} – liczba szkół zasadniczych zawodowych, techników i liceów profilowanych z dostępem do Internetu w ogólnej liczbie szkół zasadniczych zawodowych, techników i liceów profilowanych w %, X_{35} – osoby niepełnosprawne poniżej 16 roku życia na 1000 ludności poniżej 16 roku życia, X_{37} – drogi publiczne powiatowe o twardej nawierzchni w km na 1 km² powierzchni powiatów. W kolejnym kroku określono typ cech diagnostycznych i do destymulant zaliczono: X_{12} , X_{15} , X_{35} , które zostały zamienione na stymulanty.

Tabela 1. Wartości statystyk opisowych charakteryzujących cechy diagnostyczne

Parametry	Cechy						
	X_3	X_5	X_{12}	X_{15}	X_{29}	X_{35}	X_{37}
	klasyczne						
Średnia	74,505	19,243	4,948	251,221	92,629	33,089	0,348
Odchylenie standardowe	29,345	18,150	2,615	131,049	9,676	9,462	0,145
Współczynnik zmienności	39,386	94,321	52,851	52,165	10,446	28,597	41,571
Skośność	-1,311	0,811	0,030	0,132	-1,751	-2,342	2,002
	pozycyjne						
Wektor medianowy Webera	80,362	18,594	4,781	236,823	93,777	33,515	0,331
Mad	82,900	14,100	4,781	237,451	96,094	37,030	0,310
Współczynnik zmienności	103,158	75,831	100,000	100,265	102,470	110,488	93,509

Źródło: obliczenia własne.

W tabeli 1 przedstawiono wartości statystyk opisowych charakteryzujących cechy diagnostyczne. Analizując uzyskane wyniki, można zauważyć, że różnice między średnimi arytmetycznymi a medianami są raczej nieduże (wyjątek stanowią cechy X_3 i X_{15}), większość cech charakteryzuje się dość silną asymetrią lewostronną (cechy X_{29} , X_{35}) oraz prawostronną – cecha X_{37} . Cechy wykazują wysoki stopień zmienności, a wartości bezwzględnego odchylenia medianowego (mad) są dla większości cech wyższe w porównaniu z wartościami odchylenia standardowego. Także wartości współczynników zmienności w podejściu pozycyjnym przyjmują wyższe wartości w porównaniu z klasycznym współczynnikiem zmienności.

3. Wyniki badania

W tabeli 2 zaprezentowano uporządkowania i klasyfikacje powiatów województwa zachodniopomorskiego pod względem obiektywnej jakości życia mieszkańców.

Tabela 2. Uporządkowania i klasyfikacje powiatów województwa zachodniopomorskiego pod względem obiektywnej jakości życia mieszkańców

Powiaty	Miernik klasyczny			Miernik pozycyjny		
	rangi	wartość miernika	klasa typologiczna	rangi	wartość miernika	klasa typologiczna
Białogardzki	15	0,2568	III	2	0,9344	II
Drawski	19	0,1264	III	3	0,9101	II
Kołobrzeski	3	0,4557	II	14	0,5188	III
Koszaliński	5	0,3486	II	15	0,5048	III
Sławiński	11	0,2934	II	4	0,8378	II
Szczeciński	13	0,2735	III	12	0,5219	III
Świdwiński	16	0,1978	III	18	0,2504	III
Wałecki	17	0,1745	III	6	0,8276	II
Koszalin	1	0,5440	I	7	0,8177	II
Choszczeński	18	0,1589	III	11	0,5911	II
Gryfiński	21	0,0391	II	21	-0,0026	IV
Łobeski	9	0,3010	II	13	0,5188	III
Myśliborski	6	0,3314	II	9	0,7386	II
Pyrzycki	7	0,3113	II	1	0,9998	II
Stargardzki	10	0,2961	II	16	0,4685	III
Szczecin	2	0,4578	II	5	0,8363	II
Goleniowski	14	0,2669	III	19	0,1481	III
Gryfiński	8	0,3089	IV	10	0,7176	II
Kamiński	20	0,0724	III	20	0,1220	III
Policki	4	0,3822	II	8	0,8070	II
Świnoujście	12	0,2856	II	17	0,3285	III

Źródło: obliczenia własne.

Badając zgodność wyników liniowego porządkowania powiatów według klasycznego i pozycyjnego miernika, przyjmując za podstawę rangi powiatów, otrzymano wartość współczynnika korelacji τ Kendalla, wynoszącą 0,257. Niska wartość

współczynnika wynika z rozbieżności w pozycjach zajmowanych przez powiaty, np. powiat białogardzki w porządkowaniu według klasycznego miernika zajmuje miejsce 15, natomiast według pozycyjnego miernika – miejsce 2. Takie duże różnice można zauważyć także w przypadku takich powiatów, jak: drawski, wałecki, kołobrzeski. Zbliżone miejsca w obu porządkowaniach zajmują powiaty: szczecinecki, świdwiński, Szczecin, gryfiński. Jednakową pozycję w obu porządkowaniach zajmują jedynie dwa powiaty: gryficki i kamieński. Zbiór powiatów został podzielony na klasy obiektywnej jakości życia, według klasycznego i pozycyjnego miernika (tab. 3).

Tabela 3. Granice skupień oraz procent powiatów należących do skupienia

Skupienia	Granice skupień wyznaczone metodą					
	klasyczną		procent powiatów należących do skupienia	pozycyjną		procent powiatów należących do skupienia
	dolna	górna		dolna	górna	
I	0,5294	X	4,8	1,1822	X	0
II	0,2801	0,5294	52,4	0,5911	1,1822	52,4
III	0,0308	0,2801	38,1	0,0000	0,5911	42,9
IV	X	0,0308	4,8	X	0,0000	4,8

Źródło: obliczenia własne.

Najkorzystniejsza sytuacja pod względem obiektywnej jakości życia mieszkańców występuje w powiecie grodzkim Koszalin; związane to jest przede wszystkim z dobrą sytuacją w zakresie ochrony środowiska. Powiaty drugiej grupy koncentrują się w północnej, nadmorskiej części województwa, wzdłuż zachodniej granicy, a dwa powiaty leżą w części środkowej (stargardzki i łobeski). Powiat gryficki został zaliczony do ostatniej czwartej grupy. Taka niekorzystna sytuacja wynika przede wszystkim ze słabych wartości z zakresu ochrony środowiska oraz dużej liczby zgonów z powodu chorób układu oddechowego na 1000 ludności. Grupę trzecią stanowią powiaty położone w części wschodniej województwa oraz powiaty kamieński i goleniowski. Zasadnicza różnica między grupowaniami polega na tym, iż w podziale uzyskanym za pomocą metody pozycyjnej uzyskano pustą pierwszą klasę. Skuteczność grupowań zweryfikowano, wyznaczając wartości wskaźników homogeniczności, heterogeniczności i poprawności skupień (tab. 4).

Tabela 4. Wartości wskaźników homogeniczności, heterogeniczności i poprawności skupień

Wskaźniki	Wariant klasyczny	Wariant pozycyjny
Homogeniczności skupień	28,934	73,605
Heterogeniczności skupień	124,573	94,779
Poprawności skupień	0,232	0,776

Źródło: obliczenia własne.

Analizując wyniki dotyczące efektywności grupowań przedstawione w tab. 4, można stwierdzić, że klasyfikacja otrzymana metodą pozycyjną charakteryzowała się mniejszą poprawnością skupień. Taka sytuacja wynikała z faktu, iż wariant pozycyjny wykazywał pewne niedoskonałości, polegające na tym, że w grupowaniu otrzymano puste skupienie pierwsze oraz jednoelementowe skupienie czwarte, a pozostałe powiaty zostały zaliczone do grupy drugiej (ponad 52% powiatów) lub trzeciej (43% powiatów).

4. Podsumowanie

W pracy przedstawiono zastosowanie klasycznego i pozycyjnego podejścia w problematyce porządkowania liniowego obiektów. Podejście pozycyjne wykorzystywało medianę przestrzenną Webera. Badanie miało charakter lokalny i dotyczyło mieszkańców powiatów województwa zachodniopomorskiego w roku 2008. Zaprezentowane ujęcie pozycyjne ma zastosowanie szczególnie wtedy, gdy w zbiorze cech występują takie cechy, które charakteryzują się silną asymetrią. Otrzymane porządki różnią się między sobą, niewiele powiatów zajmuje zbliżone pozycje w porządkowaniach, a tylko dwa obiekty (gryficki i kamieński) były na tych samych miejscach. Duże rozbieżności w rankingach potwierdza wartość współczynnika korelacji τ Kendalla, która wyniosła 0,257. Na podstawie otrzymanych porządkowań utworzono grupy typologiczne powiatów. W podejściu klasycznym otrzymano cztery klasy, jednak do pierwszej grupy należał tylko jeden obiekt. W podejściu pozycyjnym uzyskano trzy klasy typologiczne. Rozkład powiatów w obu grupowaniach był zbliżony, najwięcej, nieco ponad 52% obiektów, należało do klasy drugiej.

Badając efektywność otrzymanych podziałów, można stwierdzić, że grupowanie otrzymane za pomocą metody klasycznej dało lepsze wyniki pod względem homogeniczności, heterogeniczności i poprawności grupowania.

Literatura

- Borys T., *Propozycja siedmiu typologii jakości życia*, „Gospodarka a Środowisko” 2008, nr 9.
- Kolman R., *Kwalitologia. Wiedza o różnych dziedzinach jakości*, Placet, Warszawa 2009.
- Lira J., Wagner W., Wysoki F., *Mediana w zagadnieniach porządkowania obiektów wielocechowych*, [w:] *Statystyka regionalna w służbie samorządu lokalnego i biznesu*, J. Paradysz (red.), Internetowa Oficyna Wydawnicza Centrum Statystyki Regionalnej, Wydawnictwo AE, Poznań 2002.
- Młodak A., *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa 2006.
- Nowak E., *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa 1991.
- Ostasiewicz W. (red.), *Ocena i analiza jakości życia*, Wydawnictwo AE, Wrocław 2004.
- Ślaby T., *Konsumpcja. Eseje statystyczne*, Difin, Warszawa 2006.
- Szukielój-Bieńkuńska A., Walczak T., *Statystyczny pomiar postępu społeczno-gospodarczego w zmieniającym się świecie*, „Wiadomości Statystyczne” 2011, nr 7-8.
- Warunki życia w województwie zachodniopomorskim w latach 2006-2008*, Urząd Statystyczny w Szczecinie, Informacje i Opracowania Statystyczne, Szczecin 2010.

THE COMPARISON OF THE CLASSICAL AND POSITIONAL TAXONOMIC ANALYSIS OF THE QUALITY OF LIFE DIFFERENTIATION IN ZACHODNIOPOMORSKIE VOIVODESHIP

Summary: The aim of this article is the comparison of the efficiency of two approaches in the construction of the development measure using classical and positional Weber's median vector in the research of the objective quality of life in Zachodniopomorskie voivodeship. In the analysis the indicators of the following were used: environment, demography, labor market, municipal infrastructure and flats, education, culture and tourism, health, road transport, incomes and expenses of the budget of local government unit. The analysis involved the year 2008. The classical approach gave the better results in terms of homogeneity, heterogeneity and correctness of clustering.

Keywords: Weber's median, quality of life, linear ordering.