

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,
Mirosław 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 Dębkowska , 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 Pawełczyk, 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 Szwarc , 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

Karolina Bartos

Uniwersytet Ekonomiczny we Wrocławiu

ANALIZA SKUPIEŃ WYBRANYCH PAŃSTW ZE WZGLĘDU NA STRUKTURĘ WYDATKÓW KONSUMPCYJNYCH OBYWATELI – ZASTOSOWANIE SIECI KOHONENA

Streszczenie: W artykule zaprezentowano wyniki analizy skupień 86 państw ze względu na udział różnego typu wydatków konsumpcyjnych obywateli w ich wydatkach konsumpcyjnych ogółem. W badania uwzględniono 12 typów wydatków, takich jak wydatki na: żywność i napoje bezalkoholowe, utrzymanie domu, odzież i obuwie, transport, wypoczynek i rekreację, alkohol i wyroby tytoniowe, telekomunikację, edukację, zdrowie, usługi hotelowe i catering, artykuły gospodarstwa domowego i usługi dla domu, pozostałe dobra. Do stworzenia grup wykorzystano sieci Kohonena.

Słowa kluczowe: zachowanie konsumentów, wydatki konsumpcyjne, analiza skupień, sieci Kohonena (SOM).

1. Wstęp

W artykule zostaną przedstawione wyniki analizy skupień 86 krajów ze względu na udział różnego typu wydatków konsumpcyjnych w wydatkach konsumpcyjnych ogółem. Celowo zastosowano zmienne wyrażające udział poszczególnych rodzajów wydatków, a nie wartość tych wydatków, by sprawdzić, czy kraje bogatsze i biedniejsze znajdują się w różnych skupieniach. Potwierdziłoby to tezę, że struktura wydatków konsumentów jest kształtowana w zależności od ilości posiadanych przez nich pieniędzy. Badanie pozwoli także odpowiedzieć na pytanie, ze względu na jakie cechy kraje w poszczególnych skupieniach są podobne, a co różnicuje skupienia między sobą oraz do jakich skupień należą kraje z wybranych obszarów geograficznych, takich jak: Europa, Azja, Afryka, Ameryka Łacińska.

2. Hierarchia potrzeb według A.H. Masłowa a wydatki konsumentów

Najbardziej znaną próbą klasyfikacji potrzeb jest hierarchia zaproponowana przez A.H. Masłowa. Podzielił on potrzeby według stopnia konieczności ich zaspokojenia.

Zaczynając od najważniejszych (potrzeb pierwszego rzędu), do których należą potrzeby fizjologiczne oraz bezpieczeństwa, a kończąc na potrzebach wyższego rzędu: przynależności, uznania i samorealizacji. Zgodnie z tą teorią zaspokajanie potrzeb podlega pewnym prawidłowościom – te znajdujące się wyżej w hierarchii mogą być zaspokojone dopiero po zaspokojeniu potrzeb niższego rzędu. Dzieje się tak, ponieważ im wyższe miejsce w hierarchii zajmuje dana potrzeba, tym mniejsze ma znaczenie dla biologicznego przetrwania, a więc tym bardziej może być opóźnione jej zaspokojenie. Jednak nie jest konieczne całkowite zaspokojenie potrzeb podstawowych, by zaczęły się pojawiać potrzeby wyższego rzędu [Rudnicki 2000]. Dlatego konsument, który zaspokoi swoje podstawowe potrzeby tylko na określonym poziomie (nie w pełni), już zacznie odczuwać chęć zaspokojenia potrzeb wyższego rzędu.

Sposób zaspokajania potrzeb przez rozmaitych konsumentów może się nieco różnić. Wpływa na to wiele czynników, takich jak: cechy charakteru danej osoby, poziom jej wykształcenia, przynależność do grup społecznych, tradycja i kultura na danym obszarze [Rudnicki 2000]. Istnieją rozmaite ograniczenia utrudniające możliwość zaspokajania potrzeb. Są to np.: bariery cenowe, podażowe (brak towaru na rynku), administracyjno-prawne (zakaz sprzedaży i konsumpcji pewnych dóbr), społeczne i religijne (zakaz spożywania niektórych rodzajów żywności i używek – wieprzowiny u żydów, alkoholu u islamistów). Wymienione ograniczenia mogą występować w poszczególnych krajach i społeczeństwach z różną siłą [Gajewski 1997]. Warto zauważyć, że do zaspokojenia potrzeb wyższego rzędu są wymagane lepsze warunki zewnętrzne: ekonomiczne, społeczne, polityczne.

Na podstawie teorii A.H. Maslowa można więc spodziewać się, że przeprowadzając badanie dotyczące wydatków konsumpcyjnych, konsumenci biedniejszych państw będą mieli inną strukturę wydatków niż z krajów bogatszych. W pierwszej kolejności będą oni zaspokajać potrzeby niższego rzędu, jak np. zaspokojenie głodu, zapewnienie dachu nad głową, ubrania, a dopiero później, o ile ich budżet im na to pozwoli, będą wydawać swoje pieniądze na potrzeby wyższego rzędu. Jednak ze względu na różnego typu czynniki pozadochodowe wpływające na to zjawisko możliwe jest wystąpienie pewnych drobnych rozbieżności w sposobie zaspokajania potrzeb obywateli krajów o podobnym stopniu zamożności.

3. Dane wykorzystane do badania

Do analizy zostały wykorzystane dane z bazy GMID (*Global Market Information Database*) firmy Euromonitor, która pozyskała je z narodowych urzędów statystycznych, Eurostatu i OECD. W badaniu poddano analizie 86 krajów, takich jak: Kanada, USA, Australia, Nowa Zelandia, 13 państw z Ameryki Łacińskiej, 35 z Europy, 26 krajów azjatyckich (w tym 8 z Bliskiego Wschodu) oraz 8 afrykańskich. Różne rodzaje wydatków konsumpcyjnych tych państw wyrażone w dol. *per capita* za 2010 r. zostały przekształcone tak, by wyrażały udział poszczególnych rodzajów

wydatków konsumentów w ich wydatkach konsumpcyjnych ogółem. Rodzaje analizowanych wydatków:

X_1 – wydatki na napoje alkoholowe i wyroby tytoniowe,

X_2 – wydatki na odzież i obuwie,

X_3 – wydatki na telekomunikację i usługi pocztowe,

X_4 – wydatki na edukację (nie zawiera wydatków na: książki, artykuły biurowe, koszty dojazdu itp.),

X_5 – wydatki na żywność i napoje bezalkoholowe,

X_6 – wydatki na zdrowie (usługi medyczne, leki, sprzęt medyczny),

X_7 – wydatki na usługi hotelowe i catering,

X_8 – wydatki na artykuły gospodarstwa domowego (jak np. sprzęt AGD, meble, dywany, wyposażenie kuchni, narzędzia) i usługi dla domu (np. sprzątanie, opieka nad dziećmi),

X_9 – wydatki na utrzymanie domu (czynsz, konserwacja i naprawy, prąd, gaz, woda),

X_{10} – wydatki na wypoczynek i rekreację (tu także: wydatki związane z utrzymaniem ogrodu i zwierząt domowych, wydatki na książki, prasę, artykuły biurowe, sprzęt audiowizualny i fotograficzny),

X_{11} – wydatki na transport,

X_{12} – wydatki na pozostałe dobra i usługi (związane z pielęgnacją ciała: kosmetyki, usługi fryzjerskie itp.; biżuteria, usługi finansowe, wydatki na ubezpieczenie, inne).

4. Opis badania i wyniki

Do badania zastosowano program *Statistica* wersja 9.0. Utworzono sieć Kohona o architekturze 1×3 . Do nauczania sieci wykorzystano zbiór o liczebności 70% zbioru wyjściowego oraz zdecydowano się na następujące parametry: 1000 iteracji, szybkość uczenia: początek 0,1, koniec 0,02, sąsiedztwo: początek 2, koniec 0.

Trzy skupienia zawierały: 34 kraje (I skupienie), 21 krajów (II skupienie) i 31 krajów (III skupienie). Tabela 1 przedstawia zestawienie wszystkich analizowanych krajów w trzech skupieniach posortowanych malejąco według sumy wydatków konsumpcyjnych poszczególnych państw *per capita* za 2010 r. wyrażonej w dol. Na podstawie tab. 1 można stwierdzić, że w pierwszej grupie są głównie kraje najbogatsze (o największych wydatkach konsumpcyjnych na osobę), w drugiej w większości średniozamożne, a w trzeciej znajdują się prawie wszystkie najbiedniejsze analizowane państwa.

Struktura wydatków konsumpcyjnych większości krajów pierwszego skupienia charakteryzuje się największym udziałem wydatków na opłaty związane z utrzymaniem domu (prąd, gaz, czynsz, woda), który wynosi średnio 21% wydatków konsumpcyjnych, oraz najmniejszym udziałem wydatków na edukację (1%). Spośród wszystkich grup kraje z I skupienia mają największy udział wydatków na rekreację

i wypoczynek (ponad 9%), utrzymanie domu (21%), pozostałe dobra i usługi (prawie 11%), hotele i catering (blisko 8%) oraz alkohol i wyroby tytoniowe (4%). Najmniejszy zaś – na żywność i napoje bezalkoholowe (14%) oraz edukację (1%). Niski udział wydatków obywateli na edukację w wydatkach konsumpcyjnych związany jest z funkcjonowaniem w większości najzamożniejszych krajów systemu edukacji finansowanego z budżetu państwa. Mimo że I grupa zawiera głównie państwa najbogatsze, to w jej strukturze znalazły się dość biedne kraje, jak Argentyna i Tajlandia, które zostały tu przydzielone ze względu na bardzo niski udział wydatków na edukację (1 i 0,8%), a wysoki na wypoczynek i rekreację (9 i 7%) oraz hotele i catering (7,5 i 7%).

Tabela 1. Zestawienie wszystkich analizowanych krajów w trzech skupieniach

I SKUPIENIE			II SKUPIENIE			III SKUPIENIE		
Nr	Kraj	Suma wydatków konsumpcyjnych per capita, USD, za 2010 rok	Nr	Kraj	Suma wydatków konsumpcyjnych per capita, USD, za 2010 rok	Nr	Kraj	Suma wydatków konsumpcyjnych per capita, USD, za 2010 rok
1	Switzerland	38310	1	United Arab Emirates	24437	1	Kuwait	13203
2	Norway	33390	2	Israel	16342	2	Chile	6442
3	USA	32851	3	Qatar	11986	3	Saudi Arabia	5836
4	Australia	31420	4	Taiwan	10926	4	Montenegro	5822
5	Denmark	26770	5	South Korea	10587	5	Russia	5414
6	Canada	25786	6	Uruguay	8416	6	Romania	4536
7	Austria	24988	7	Turkey	7588	7	Bosnia-Herzegovina	3932
8	Japan	24617	8	Brazil	6620	8	Serbia	3782
9	France	23754	9	Bahrain	6242	9	Macedonia	3626
10	Sweden	23615	10	Mexico	6172	10	Kazakhstan	3329
11	Finland	22998	11	Venezuela	5820	11	Belarus	3158
12	United Kingdom	22514	12	Costa Rica	4792	12	Jordan	2915
13	Germany	21954	13	Bulgaria	4229	13	Iran	2690
14	Belgium	21774	14	South Africa	4141	14	Guatemala	2498
15	Ireland	21753	15	Colombia	3942	15	Tunisia	2342
16	Netherlands	21205	16	Malaysia	3741	16	Azerbaijan	2252
17	Greece	21027	17	Dominican Republic	3713	17	Egypt	2236
18	Italy	20599	18	Peru	3207	18	Georgia	2013
19	Hong Kong, China	19503	19	Ecuador	2879	19	Turkmenistan	2011
20	New Zealand	19120	20	China	1514	20	Morocco	1894
21	Spain	18051	21	Bolivia	1179	21	Ukraine	1848
22	Singapore	16312				22	Indonesia	1735
23	Portugal	14404				23	Algeria	1517
24	Slovenia	13938				24	Philippines	1448
25	Czech Republic	9589				25	India	808
26	Slovakia	9441				26	Vietnam	783
27	Croatia	9255				27	Cameroon	768
28	Poland	7490				28	Pakistan	762
29	Estonia	7247				29	Uzbekistan	708
30	Hungary	6967				30	Nigeria	659
31	Lithuania	6901				31	Kenya	628
32	Latvia	6207						
33	Argentina	4804						
34	Thailand	2513						

Źródło : opracowanie własne.

Większość obywateli państw drugiej grupy przeznaczają najwięcej swoich pieniędzy na żywność i napoje bezalkoholowe (21%) oraz na utrzymanie domu (prawie 19%), najmniej zaś na alkohol i wyroby tytoniowe (2%). W porównaniu z innymi skupieniami państwa drugiej grupy mają najwyższy udział wydatków na edukację

(4,7%) oraz na telekomunikację i usługi pocztowe (5%), a najniższy na alkohol i wyroby tytoniowe (2%). Drugie skupienie zawiera głównie kraje średniozamożne, wyjątkiem są tutaj Zjednoczone Emiraty Arabskie, Izrael, Katar, Tajwan, Korea Południowa. Zjednoczone Emiraty Arabskie znalazły się w tej grupie głównie ze względu na bardzo niski udział wydatków na alkohol i wyroby tytoniowe (0, 2%) oraz wypoczynek i rekreację (2%), niewielki na hotele i catering (4,8%), a wysoki na telekomunikację i usługi pocztowe (7%). Izrael, Katar, Tajwan i Korea Południowa mają zbyt wysoki udział wydatków na edukację (odpowiednio 6,5, 7, 6,6 i 6,3%), by znaleźć się w I skupieniu, oraz zbyt mały na alkohol i wyroby tytoniowe (1,6, 0,2, 1,9, 2,5%). Tak więc mimo że niektóre państwa należą do krajów najzamożniejszych, ich struktura wydatków różni się od struktury wydatków krajów I skupienia. Dzieje się tak, ponieważ na sposób wydatkowania pieniędzy wpływają nie tylko dochody konsumentów, ale także wiele innych czynników, jak polityka państwa (bezpłatny system edukacji) czy czynniki religijne (np. niskie spożycie alkoholu w krajach muzułmańskich).

Struktura wydatków konsumpcyjnych większości krajów trzeciej grupy charakteryzuje się największym udziałem wydatków na żywność i napoje bezalkoholowe (średnio aż 35%, ale są państwa z udziałem nawet powyżej 45%, np. Kamerun i Azerbejdżan). W porównaniu z pozostałymi ta grupa ma najniższy udział wydatków na: wypoczynek i rekreację (3%); hotele i catering (prawie 4%); transport (9%); inne dobra i usługi (6%) oraz na utrzymanie domu (16%), a najwyższy na żywność i napoje bezalkoholowe oraz odzież i obuwiu (6,6%). W przypadku Kuwejtu trudno jest wskazać jedną główną przyczynę zaklasyfikowania tego kraju do III skupienia. Występuje za to wiele mniejszych, do których należą: niski udział wydatków w porównaniu z krajami ze skupienia I i II na: hotele i catering (2,7%), inne dobra i usługi (6,7%), transport (8%), wypoczynek i rekreację (4%), a zbyt wysoki na odzież i obuwiu (8,8%). Ponadto wystąpił zbyt duży udział wydatków na edukację (3%) i telekomunikację (4,5%) w porównaniu z I skupieniem, a zbyt mały, by Kuwejt został przydzielony do skupienia II.

Tabela 2 obrazuje, do której grupy trafiły państwa poszczególnych obszarów geograficznych.

Pierwsza grupa to w większości kraje europejskie, bogatsze kraje Azji oraz wspomniane wcześniej wyjątki Tajlandia i Argentyna. Nieujęte w tabeli kraje, takie jak: USA, Kanada, Australia i Nowa Zelandia, również znalazły się w I skupieniu. Prawie 50% krajów drugiej grupy pochodzi z Ameryki Łacińskiej, 8 z Azji (głównie z Bliskiego Wschodu), 2 z Europy (Turcja i Bułgaria) i 1 z Afryki (najbogatszy afrykański kraj – RPA). Trzecie skupienie zawiera prawie wszystkie analizowane afrykańskie państwa, większość azjatyckich, 2 z Ameryki Łacińskiej oraz – co mogłoby dziwić – aż 8 europejskich. Związane jest to z bardzo dużym udziałem wydatków na żywność i napoje bezalkoholowe w tych krajach (na Ukrainie i Białorusi wynosi on aż 42%, w Rosji, Macedonii, Rumunii ok. 30%), a dość niskim na utrzymanie domu (na Ukrainie tylko 9%).

Tabela 2. Kraje poszczególnych obszarów geograficznych a numer skupienia

Obszar geograficzny	Numer skupienia		
	I	II	III
Europa	25	2 (Turcja, Bułgaria)	8 (Rosja, Czarnogóra, Ukraina, Bośnia i Hercegowina, Macedonia, Rumunia, Serbia, Białoruś)
Azja	4 (Japonia, Hongkong, Singapur, Tajlandia)	8 (Zjednoczone Emiraty Arabskie, Katar, Izrael, Tajwan, Korea Południowa, Bahrajn, Malezja, Chiny)	14
Ameryka Łacińska	1 (Argentyna)	10	2 (Chile, Gwatemala)
Afryka	0	1 (RPA)	7 (Tunezja, Egipt, Maroko, Algieria, Kamerun, Nigeria, Kenia)

Źródło: opracowanie własne.

Tabela 3. Średni udział poszczególnych wydatków w powstałych skupieniach

Skupienie	Alkohol i tytoń	Odzież i obuwie	Telekomunikacja i poczta	Edukacja	Żywność i napoje bezalkoholowe	Zdrowie	Hotele i catering	A GD	Utrzymanie domu	Wypoczynek	Transport	Inne
I	0,04096	0,05025	0,02822	0,01384	0,14499	0,04933	0,07766	0,05847	0,20968	0,09413	0,12556	0,10691
II	0,02293	0,04615	0,05133	0,04685	0,21242	0,04837	0,06448	0,05510	0,18802	0,05087	0,12522	0,08824
III	0,02847	0,06641	0,03351	0,02824	0,34636	0,04309	0,03897	0,06525	0,16078	0,03433	0,09388	0,06070

Źródło: opracowanie własne.

Średni udział wydatków na żywność i napoje bezalkoholowe w ogólnych wydatkach konsumpcyjnych zmniejsza się w zależności od zamożności konsumentów (tab. 3). W I skupieniu (gdzie występują w większości najbogatsze kraje) jest najmniejszy, a w III skupieniu (w większości najbiedniejsze kraje) zwiększa się aż do 35%. Odwrotną wyraźną zależność można zaobserwować w przypadku wydatków na utrzymanie domu oraz wypoczynek i rekreację. Tam udział wydatków wzrasta wraz z zamożnością grupy państw.

W III skupieniu struktura wydatków jest najbardziej nierównomiernie rozłożona (charakteryzuje się największą wariancją). Konsumenti biedniejszych krajów przeznaczają średnio aż 34,6% swoich wydatków na żywność i napoje bezalkoholowe, 16% na utrzymanie domu oraz 6,6% na odzież i obuwie, a więc na zaspokojenie tylko swoich potrzeb niższego rzędu przeznaczają aż 57,2% wydatków. W II skupieniu jest to 44,6%, a w I skupieniu 40%. Odwrotna sytuacja występuje w przypadku zaspokajania potrzeb wyższego rzędu (wydatki na usługi hotelowe i catering, wydatki

na wypoczynek i rekreację, wydatki na pozostałe dobra i usługi). Tutaj to obywatele krajów I skupienia wydają stosunkowo najwięcej swoich pieniędzy (prawie 28%) w porównaniu z innymi skupieniami (II skupienie – ponad 20%, a III skupienie tylko 13%).

5. Podsumowanie

W badaniu wykorzystano dane dotyczące udziału poszczególnych rodzajów wydatków (a więc nie określono ich wartości), a mimo to powstały skupienia zawierające: przeważnie kraje bogate (I skupienie), w większości średniozamożne (II skupienie) oraz głównie biedniejsze (III skupienie). Mogłoby to oznaczać, że konsumenci posiadający porównywalną ilość środków pieniężnych mają podobną do siebie strukturę wydatków konsumpcyjnych. Jednak nie można tego jednoznacznie stwierdzić ze względu na pojawienie się w wynikach badania wyjątków (np. Zjednoczone Emiraty Arabskie, Katar, Argentyna, Tajlandia, Kuwejt). Występujące wyjątki potwierdziły natomiast tezę, że na strukturę wydatków nie wpływają jedynie czynniki dochodowe, ale ważne są również inne, np. polityka państwa czy szczególnie istotne wśród krajów muzułmańskich czynniki społeczne i religijne.

Potwierdziła się także teoria A.H. Masłowa, że w pierwszej kolejności konsumenci zaspokajają potrzeby niższego rzędu, jak np. zaspokojenie głodu, zapewnienie dachu nad głową, ubrania, a dopiero później wydatkują swoje pieniądze na potrzeby wyższego rzędu. Dlatego najbiedniejsze kraje trzeciego skupienia przeznaczają na zaspokojenie potrzeb niższego rzędu najwięcej swoich wydatków, bo aż ponad 57%, a obywatele najbogatszych państw przeznaczają najwięcej w stosunku do krajów z innych skupień na potrzeby wyższego rzędu.

Literatura

- Gajewski S., *Zachowanie się konsumenta a współczesny marketing*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 1997.
- Larose D.T., *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Lula P., *Wybrane możliwości zastosowań sieci neuronowych i programu STATISTICA Neural Networks* [w:] *Sieci Neuronowe*, StatSoft, Kraków 1999.
- Lula P., Tadeusiewicz R., Paliwoda-Pękosz G., *Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu*, Wydawnictwo AE, Kraków 2007.
- Rudnicki L., *Zachowanie konsumentów na rynku*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2000.

CLUSTER ANALYSIS OF SELECTED COUNTRIES DUE TO THE STRUCTURE OF THEIR CITIZENS' CONSUMER EXPENDITURES – THE USE OF KOHONEN NETWORKS

Summary: The article presents the results of cluster analysis of 86 countries due to the ratio of various types of citizens' consumer expenditures in their total consumer expenditures. The study included 12 types of expenditures such as expenditure on: food and non-alcoholic beverages, housing, clothing and footwear, transport, leisure and recreation, alcoholic beverages and tobacco, communication, education, health goods and medical services, hotels and catering, household goods and services, miscellaneous goods and services. Kohonen network was used to create the clusters.

Keywords: consumer behavior, consumer expenditures, cluster analysis, Kohonen networks (SOM).