

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

Katarzyna Cheba

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

Joanna Hołub-Iwan

Uniwersytet Szczeciński

WYKORZYSTANIE ANALIZY KORESPONDENCJI W SEGMENTACJI RYNKU USŁUG MEDYCZNYCH

Streszczenie: Zachowania nabywczce (rynkowe) pacjentów warunkowane są bezpośrednio dostępem do usług medycznych, wynikającym z przyjętych w danym kraju rozwiązań systemowych. W pracy do segmentacji rynku usług medycznych wykorzystano metody wielowymiarowej analizy porównawczej: analizę korespondencji oraz metodę Warda. Analiza została przeprowadzona na podstawie danych opisujących zachowania zdrowotne ponad 1000 mieszkańców Polski.

Słowa kluczowe: zachowania nabywczce (rynkowe) pacjentów, analiza korespondencji, metoda Warda.

1. Wstęp

Pojęcie zachowań nabywczych (rynkowych) konsumentów, szeroko opisywane w literaturze z zakresu marketingu [Jachnis 2007; Światowy 2006; Rudnicki 2000; Gajewski 1997], dopiero od niedawna zaczęło funkcjonować na gruncie dyscyplin zajmujących się zdrowiem publicznym [Hołub-Iwan 2013]. Wynika to z wielu przyczyn, z których najważniejsze dotyczą specyfiki rynku usługi medycznych i samej usługi medycznej, jak również z faktu, iż dopiero od niedawna w literaturze przedmiotu pojawiły się informacje na temat wpływu na zdrowie i zachowania pacjentów uwarunkowań związanych z funkcjonowaniem otoczenia organizacyjnego i instytucjonalnego ochrony zdrowia. Uznanie wpływu na zdrowie organizacji systemu opieki zdrowotnej nie jest w literaturze z zakresu zdrowia publicznego jednoznaczne, np. Światowa Organizacja Zdrowia (WHO) wśród uwarunkowań zdrowia nie wymienia organizacji opieki zdrowotnej. Koncepcja wyodrębnienia zachowań nabywczych pacjentów jako obszaru badań i analiz zainspirowana została teoriami naukowymi uznającymi wpływ na zdrowie czynników związanych z organizacją ochrony zdrowia, takimi jak: pola zdrowia Lalonda, tzw. holistyczny

model zdrowia publicznego, model Bluma, mandala zdrowia [Lalonde 1974]. Propozycja wyodrębnienia zachowań nabywczych pacjentów wiąże się z szeroko pojętymi warunkami dostępu społeczeństwa do usług medycznych, w tym oczywiście z przyjętymi w danym kraju rozwiązaniami systemowymi. Warunki te stanowią otoczenie, w którym funkcjonują pacjenci. Otoczenie, choć faktycznie dotyczy aspektów instytucjonalnych, ustawowych, stanowi też (w jakimś sensie) środowisko, w którym funkcjonują pacjenci.

Celem pracy jest segmentacja pacjentów z uwzględnieniem czynników związanych z organizacją systemu ochrony z zdrowia w Polsce i poszukiwanie związków pomiędzy zmiennymi demograficzno-społecznymi konsumentów a zmiennymi opisującymi ich zachowania w procesie podejmowania decyzji rynkowych na quasi-ryнку usług medycznych. Do realizacji celu badawczego posłużono się wielowymiarową analizą korespondencji. Badanie zachowań nabywczych pacjentów jest częścią projektu badawczego finansowanego ze środków na naukę w latach 2010-2013, pt.: „Analiza zachowań pacjentów i ich wpływ na funkcjonowanie systemu opieki zdrowotnej w Polsce”¹.

2. Metoda badawcza

Do segmentacji pacjentów na rynku usług medycznych wykorzystano analizę korespondencji, której celem jest identyfikowanie jednoczesnego współwystępowania poszczególnych kategorii rozpatrywanych zmiennych nominalnych [Andersen 1997]. Analiza korespondencji jako rodzaj analizy czynnikowej pozwala na identyfikację związków pomiędzy zmiennymi oraz obiektami głównie w formie graficznej [Panek 2009].

W pracy obliczenia do wielowymiarowej analizy korespondencji przeprowadzono na symetrycznej macierzy blokowej macierzy Burta, w której na głównej przekątnej znajdują się macierze diagonalne, zawierające liczebności poszczególnych kategorii. Bloki pozadiagonalne są tablicami kontyngencji między parami zmiennych [Stanimir 2005]. Macierz Burta (B) jest iloczynem wewnętrznym macierzy znaczników (Z^2):

$$B = Z^T Z. \quad (1)$$

Wymiar rzeczywistej przestrzeni współwystępowania (K) wyznacza się na podstawie wzoru:

$$K = \sum_{q=1}^Q (J_q - 1), \quad (2)$$

gdzie: J_q – liczba kategorii zmiennej q ($q = 1, 2, \dots, Q$), Q – liczba zmiennych.

¹ Projekt badawczy MNiSW nr N N114 134034.

² Macierz znaczników Z (układu, kodów) jest zbudowana w taki sposób, że poszczególne wiersze odpowiadają kolejnym obserwacjom, a kolumny – wariantom wszystkich zmiennych.

Przy wyborze istotnych dla badania wartości własnych ($\lambda_{B,k}$) stosuje się kryterium Greenacre'a według, którego wartości te powinny spełniać nierówność:

$$\lambda_{B,k} > \frac{1}{Q}, \quad (3)$$

gdzie: $\lambda_{B,k}$ – jest k -tą wartością własną ($k = 1, 2, \dots, K$) otrzymaną w trakcie analizy macierzy Burta, Q – liczba zmiennych.

W celu podwyższenia jakości odwzorowania przeprowadza się modyfikację wartości własnych zgodnie z propozycją Greenacre'a na podstawie wzoru [Greenacre 1984; Bąk i Wawrzyniak 2009, s. 7]:

$$\tilde{\mathbf{F}} = \mathbf{F}^* \cdot \mathbf{\Gamma}^{-1} \cdot \tilde{\mathbf{\Lambda}}, \quad (4)$$

gdzie: $\tilde{\mathbf{F}}$ – macierz zmodyfikowanych wartości współrzędnych dla kategorii badanych zmiennych o wymiarze $K \times k$, \mathbf{F}^* – macierz pierwotnych wartości współrzędnych dla kategorii badanych zmiennych o wymiarze $K \times k$, $\mathbf{\Gamma}^{-1}$ – diagonalna macierz odwrotna wartości osobliwych o wymiarze $k \times k$, $\tilde{\mathbf{\Lambda}}$ – diagonalna macierz zmodyfikowanych wartości własnych o wymiarze $k \times k$, K – wymiar rzeczywistej przestrzeni współwystępowania.

3. Materiał badawczy

W badaniach wykorzystano informacje uzyskane w trakcie badań ankietowych przeprowadzonych przez autorki w pierwszej połowie 2011 roku na próbie losowej obejmującej 1067 dorosłych Polaków (powyżej 15 roku życia), zamieszkałych na terytorium Rzeczypospolitej Polskiej, przy błędzie standardowym frakcji nieprzekraczającym 3% wartości cechy, przy poziomie ufności $\alpha = 0,05$. Próbę dobrano metodą warstwową – proporcjonalną. Wyłoniona próba badawcza jest reprezentatywna ze względu na: płeć, wiek (w przedziałach: 15-29 lat, 30-44 lata, 45-59 lat oraz 60 lat i więcej), region (klasyfikacja NUTS-1: region centralny, południowy, wschodni, północno-zachodni, południowo-zachodni oraz północny), miejsce zamieszkania (wieś, miasto do 30 tys., miasto do 200 tys., miasto powyżej 200 tys.).

Kwestionariusz badania ankietowego podzielony został na 5 części, w ramach których pozyskano informacje obejmujące m.in.: a) ocenę funkcjonowania państwowej służby zdrowia, w tym zachowania pacjentów w przypadku trudności ze skorzystaniem z usług państwowej służby zdrowia, b) ocenę funkcjonowania prywatnej służby zdrowia, w tym zachowania pacjentów obejmujące intensywność korzystania z usług prywatnych czy wielkość wydatków na leczenie w systemie prywatnym, c) identyfikację etapów procesu decyzyjnego oraz preferencji pacjentów związanych np. ze sposobami współpłacenia za usługi medyczne czy organizacją procesu obsługi pacjenta, d) wskazanie możliwości promowania zachowań prozdrowotnych pacjentów związanych np. z korzystaniem z badań profilaktycz-

nych, e) opis postaw pacjentów dotyczących m.in. odpowiedzialności za zdrowie, solidaryzmu czy równości w dostępie do badań [Hołub-Iwan 2013].

W celu weryfikacji przedstawionego w pracy celu badania analizie poddano odpowiedzi opisujące sposób korzystania pacjentów z usług medycznych: K1 – korzystanie tylko z usług państwowej służby zdrowia, K2 – korzystanie w pierwszej kolejności z usług państwowej służby zdrowia, a dopiero w sytuacji na przykład długiego oczekiwania na realizację usługi wybór usług świadczonych prywatnie, K3 – korzystanie w pierwszej kolejności z usług prywatnych. Natomiast identyfikując zmienne demograficzne, przyjęto następujące warianty odpowiedzi: a) płeć: K – kobieta, M – mężczyzna, b) miejsce zamieszkania: W – wieś, M1 – miasto do 30 tys., M2 – miasto 30-200 tys., M3 – miasto powyżej 200 tys. c) wykształcenie: P – podstawowe, Z – zawodowe, S – średnie, WY – wyższe, d) typ rodziny: BD – osoby i małżeństwa bez dzieci, D1 – rodziny z dziećmi do 15 lat, D2 – rodziny z dziećmi powyżej 15 lat, PG – puste gniazdo. Dodatkowo zidentyfikowano również zmienne charakteryzujące postawy pacjentów związane z korzystaniem ze służby zdrowia: a) zachowania pacjentów związane z wyborem sposobu korzystania z usług służby zdrowia: L1 – wybór sposobu korzystania uzależniony od opinii na temat lekarza, L2 – wybór uzależniony od opinii na temat placówki medycznej, L3 – inne czynniki decydujące o wyborze, c) zachowania pacjentów związane z korzystaniem z badań profilaktycznych: B1 – wykonywanie badań profilaktycznych jedynie po uzyskaniu skierowania (bezpłatnie), B2 – rezygnacja z wykonywania badań profilaktycznych, B3 – wykonywanie badań prywatnie.

4. Opis wyników badania

Analizę korespondencji przeprowadzono w celu wykrycia wzajemnych powiązań pomiędzy zmiennymi opisującymi sposób korzystania przez pacjentów z usług medycznych oraz zmiennymi charakteryzującymi postawy pacjentów związane z korzystaniem z tych usług a zmiennymi o charakterze demograficzno-społecznym. Utworzono w tym celu macierz Burta o wymiarach 23×23 , a wymiar rzeczywistej przestrzeni współwystępowania zidentyfikowanych odpowiedzi na pytania dla badanych siedmiu zmiennych wyniósł 16. Dzięki wykorzystaniu kryterium Greenacre'a³, sprawdzono, w jakim stopniu wartości własne o niższym wymiarze wyjaśniają inercję całkowitą ($\lambda = 2,2857$). Wyniki tego etapu badania przedstawiono w tabeli 1.

Z informacji przedstawionych w prezentowanej tabeli wynika, że istotne dla badania są inercje główne przyjmujące wartości własne co najwyżej 6. Jako kryterium wyboru przestrzeni rzutowania zastosowano kryterium łokcia, zgodnie z któ-

³ Kryterium to zakłada, że za istotne dla badania uznaje się inercje główne większe niż $\frac{1}{Q} > \frac{1}{7} > 0,1429$.

Tabela 1. Wyniki analizy korespondencji

Liczba wymiarów K	Wartości osobliwe γ_k	Wartości własne λ_k	Procent bezwładności λ_k/λ	Skumulowany procent bezwładności $\bar{\tau}_k$
1	0,4942	0,2443	10,6864	10,6864
2	0,4419	0,1952	8,5421	19,2284
3	0,4149	0,1722	7,5320	26,7604
4	0,4065	0,1653	7,2304	33,9908
5	0,3971	0,1577	6,8987	40,8895
6	0,3878	0,1504	6,5797	47,4692
7	0,3825	0,1413	6,4017	53,8710
8	0,3799	0,1403	6,3139	60,1848
9	0,3676	0,1352	5,9134	66,0982
10	0,3609	0,1302	5,6970	71,7952
11	0,3588	0,1287	5,6319	77,4271
12	0,3538	0,1252	5,4753	82,9024
13	0,3389	0,1149	5,0252	87,9275
14	0,3167	0,1003	4,3885	92,3160
15	0,3116	0,0971	4,2478	96,5638
16	0,2803	0,0785	3,4362	100,0000

Źródło: obliczenia własne.

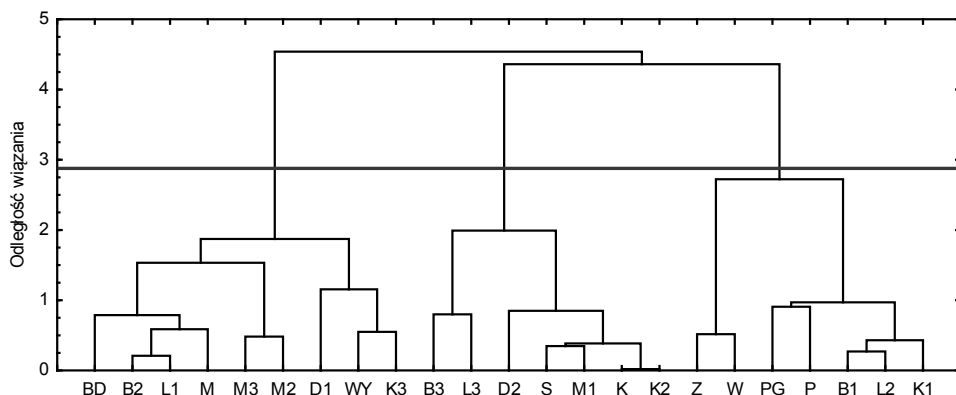
rym właściwą przestrzenią rzutowania jest przestrzeń wskazana przez numer wartości własnej, dla której nastąpił uskok na wykresie wartości własnych [Clausen 1998]. Zgodnie z tym kryterium za właściwy wymiar przestrzeni odwzorowania przyjęto przestrzeń trójwymiarową, dla której stopień wyjaśniania inercji wyniósł 26,76%. W wyniku zastosowania modyfikacji zgodnie z kryterium Greenacre'a trzy pierwsze wartości własne stanowią 59,64% zmodyfikowanej inercji całkowitej, co oznacza, że wyniku przyjęcia trzech pierwszych wymiarów udało się wyjaśnić większy niż przed modyfikacją procent inercji całkowitej. Wyniki tego etapu badania przedstawiono w tabeli 2.

Tabela 2. Wyniki analizy korespondencji po modyfikacji zgodnie z kryterium Greenacre'a

Liczba wymiarów K	$\bar{\lambda}_k$	$\bar{\lambda}_k/\bar{\lambda}$	$\bar{\tau}_k$
1	0,1680	25,6650	25,6650
2	0,1217	18,5859	44,2509
3	0,1007	15,3869	59,6378
4	0,0946	14,4526	74,0904
5	0,0880	13,4368	87,5272
6	0,0817	25,6650	100,0000
	$\bar{\lambda}_k = 0,6548$		

Źródło: obliczenia własne.

Zgodnie z propozycją przedstawioną w pracach [Stanimir 2005, Bąk 2009] w celu przedstawienia powiązań pomiędzy wariantami rozpatrywanych zmiennych wykorzystano metodę Warda. Wyniki tego etapu badania przedstawiono na rysunku 1, linią poziomą zaznaczono etap przerwania łączenia klas.



Rys. 1. Wyniki analizy korespondencji z uwzględnieniem modyfikacji wartości własnych w połączeniu z wynikami analizy Warda

Źródło: opracowanie własne.

Na podstawie wyników analizy rozwiązania uzyskanego za pomocą wielowymiarowej analizy korespondencji w połączeniu z wynikami uzyskanymi metodą Warda można wskazać na istnienie powiązań pomiędzy kategoriami analizowanych zmiennych opisujących zachowania nabywcze (rynkowe) pacjentów a zmiennymi o charakterze społeczno-demograficznym. Otrzymane wyniki pozwoliły na wyodrębnienie następujących klas:

Klasa I (K1, B1, L2, Z, W, P, PG): obejmuje osoby z wykształceniem zawodowym lub podstawowym, mieszkające na wsi, funkcjonujące w ramach tzw. pustego gniazda (z usamodzielnionymi dziećmi), korzystające tylko z usług państwowej służby zdrowia, wykonujące badania profilaktyczne jedynie po uzyskaniu skierowania (bezpłatnie), wybierające lekarza na podstawie opinii na temat placówki medycznej.

Klasa II (K2, B3, L3, K, S, D2, M1): obejmuje kobiety, mieszkanki małych miast, z wykształceniem średnim, posiadające dzieci w wieku powyżej 15 lat, korzystające w pierwszej kolejności z usług państwowej służby zdrowia, a dopiero w sytuacji na przykład długiego oczekiwania na realizację usługi dokonujące wyboru usług świadczonych prywatnie, uwzględniające przy wyborze sposobu korzystania z usług medycznych inne czynniki niż opinia na temat lekarza czy placówki medycznej, wykonujące badania profilaktyczne prywatnie.

Klasa III (K3, L1, B2, M, BD, D1, WY, M2, M3): obejmuje mężczyzn z wykształceniem wyższym, bez dzieci lub z dziećmi w wieku do 15 lat, mieszkających

w miastach powyżej 30 tys. mieszkańców, korzystających w pierwszej kolejności z usług prywatnych, uzależniających wybór lekarza od opinii na jego temat oraz niekorzystających z badań profilaktycznych.

5. Wnioski

Celem wykorzystania przedstawionej w pracy metody wielowymiarowej analizy korespondencji była identyfikacja powiązań pomiędzy kategoriami cech wyrażonych na skali nominalnej, opisujących zachowania zdrowotne (rynkowe) pacjentów, a zmiennymi o charakterze społeczno-demograficznym. Identyfikacja powiązań pomiędzy kategoriami opisywanych zmiennych pozwoliła na identyfikację przynależności konsumentów do poszczególnych segmentów. Istotą segmentacji rynku jest identyfikacja w ramach wyodrębnionej grupy nabywców homogenicznych podgrup – segmentów rynku. Z punktu widzenia marketingu sens takiej analizy związany jest bezpośrednio z możliwością łatwiejszego dotarcia do poszczególnych segmentów z przekazami promocyjnymi, ofertą produktów. Przeprowadzone badania pozwoliły na wyodrębnienie trzech segmentów pacjentów, zróżnicowanych przede wszystkim ze względu na sposób korzystania z usług medycznych. Widać wyraźny podział na pacjentów korzystających jedynie z państwowej służby zdrowia (klasa I), pacjentów o zróżnicowanych preferencjach, uzależnionych głównie od czasu oczekiwania na realizację usługi (klasa II), oraz pacjentów korzystających w pierwszej kolejności z usług prywatnych (klasa III). Zaproponowany podział jest charakterystyczny dla tego rynku usług, na którym czynniki takie, jak płeć, wiek, wykształcenie, miejsce zamieszkania oraz typ rodziny, warunkują zachowania nabywcze pacjentów.

Literatura

- Andersen E.B. (1997), *Introduction to the statistical analysis of categorical data*, Springer – Verlag, Berlin.
- Bąk I., Wawrzyniak K. (2009), *Badanie jakości środowiska naturalnego w województwie zachodniopomorskim z wykorzystaniem wielowymiarowej analizy korespondencji*, „Oeconomica” 275 (57), s. 5-15.
- Clausen S.E. (1998), *Applied correspondence analysis. An introduction*, Sage, Univeristy Paper 121, Newbury Park.
- Gajewski S. (1997), *Zachowanie się konsumenta a współczesny marketing*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Greenacre M, *Theory and Applications of Correspondence Analysis*, Academic Press, London.
- Hancock T. (1993), *Health, human development and the community ecosystem: Three ecological models*, „Health Promotion International”, vol. 6, no. 1, s. 95-99.
- Hołub-Iwan J. (2013), *Koncepcja systemu informacji marketingowej w zarządzaniu zdrowiem publicznym*, PTE, Szczecin.
- Jachnis A. (2007), *Psychologia konsumenta*, Oficyna Wydawnicza Branta, Bydgoszcz – Warszawa.

- Lalonde M. (1974), *A New Perspective on the Health of Canadians, A Working Document of Canada*, Ottawa.
- Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, SGH, Warszawa.
- Rudnicki L. (2000), *Zachowanie konsumentów na rynku*, PWE, Warszawa.
- Stanimir A. (2005), *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wyd. AE, Wrocław.
- Światowy G. (2006), *Zachowania konsumentów. Determinanty oraz metody poznania i kształtowania*, PWE, Warszawa.

THE APPLICATION OF THE CORRESPONDENCE ANALYSIS OF PATIENTS SEGMENTATION ON THE MEDICAL SERVICE MARKET

Summary: Purchasing behavior (market behavior) of patients is conditioned directly by the access to the medical services, resulting from national system solutions. In the paper, in order to conduct the medical service market segmentation, there were applied methods of multidimensional comparative analysis: correspondence analysis and Ward's method. The analysis was based on data describing the health behavior of more than 1,000 Polish citizens.

Keywords: purchasing behavior (market behavior) of the patients, correspondence analysis, Ward's method.