

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

Adam Depta

Politechnika Łódzka

ZASTOSOWANIE ANALIZY KORESPONDENCJI DO OCENY JAKOŚCI ŻYCIA LUDNOŚCI NA PODSTAWIE KWESTIONARIUSZA SF-36V2

Streszczenie: W pracy podjęto próbę aplikacji analizy korespondencji do oceny jakości życia osób poddanych badaniom za pomocą kwestionariusza jakości życia SF-36v2™ Health Survey, w ośmiu podskalach: sprawność fizyczna, ograniczenia aktywności z powodu zdrowia fizycznego, dolegliwości bólowe, ogólna percepcja zdrowia, witalność, funkcjonowanie społeczne, zdrowie psychiczne oraz ograniczenie aktywności z powodu problemów emocjonalnych. Dane dotyczą lat 2011-2013.

Słowa kluczowe: jakość życia, analiza korespondencji, kwestionariusz jakości życia SF-36v2™.

1. Wstęp

W pracy podjęto próbę aplikacji analizy korespondencji do oceny jakości życia osób badanych za pomocą kwestionariusza jakości życia SF-36v2.

Jąkanie jest nie płynnością mówienia, spowodowaną nadmiernymi skurczami mięśni oddechowych, fonacyjnych lub artykulacyjnych. Niepłynności tej towarzyszą różnorodne reakcje indywidualne lub społeczne, zakłócające komunikację międzyludzką [Tarkowski 2001]. Jako wada wymowy obniża jakość życia, która jest stopniem satysfakcji człowieka (społeczeństwa) z całej swojej egzystencji, jest więc sumą indywidualnego lub zbiorowego odczucia istniejących warunków, a zarazem ich oceną. Wada ta często rozpatrywana jest w sensie medycznym.

Według Światowej Organizacji Zdrowia (WHOQOL Group – World Health Organization Quality of Life Group) jakość życia jest określana jako „sposób postrzegania przez jednostki swoich pozycji w życiu w kontekście kultury i systemu wartości, w których egzystują w powiązaniu z własnymi celami, oczekiwaniami, standardami i obawami; jest to szeroko pojęta koncepcja, na którą w sposób kompleksowy wpływają zdrowie fizyczne jednostki, jej stan psychiczny, relacje społeczne, stopień niezależności oraz jej stosunek do znaczących cech otaczającego środowiska” [WHOQOL Group 1993].

Analiza korespondencji pozwala na opisanie i ocenę jakości życia, jest jedną z wyspecjalizowanych metod eksploracji danych. Ma szeroki zakres zastosowań oraz możliwości graficznej prezentacji wyników badań. Wyniki otrzymane przy jej wykorzystaniu są przejrzyste, co bardzo ułatwia ich interpretację. Ponadto analiza ta umożliwia rozwiązanie jednego z najtrudniejszych zadań, którym jest trafne rozpoznanie współwystępowania kategorii zmiennych lub obiektów zmierzonych na skali nominalnej i porządkowej [Greenacre 1984; Stanimir 2005; Walesiak, Gatnar (red.) 2009].

Zasadniczym celem pracy jest zatem identyfikacja struktury powiązań pomiędzy kategoriami zmiennych określających wyniki w podskalach, takich jak: sprawność fizyczna (*physical functioning* – F), ograniczenia aktywności z powodu zdrowia fizycznego (*role limitations due to physical problems* – R), dolegliwości bólowe (*bodily pain* – P), ogólna percepcja zdrowia (*general health perception* – H), witalność (*vitality* – V), funkcjonowanie społeczne (*social functioning* – S), zdrowie psychiczne (*mental health* – W) oraz ograniczenie aktywności z powodu problemów emocjonalnych (*role limitation due to emotional problems* – E), a kategorią zmienną o charakterze społeczno-demograficznym, czyli grupą wiekową. Ponadto w artykule zaprezentowano w przestrzeni dwuwymiarowej oryginalne konfiguracje punktów reprezentujących podskale i grupy wiekowe przy zachowaniu pełnej lub prawie pełnej informacji o zróżnicowaniu wierszy i kolumn tabeli wielodzielczej.

Analizowane zmienne odnoszą się do różnych pacjentów jękających się, którzy odpowiadali na pytania w kwestionariuszu SF-36v2 w dniu zgłoszenia się na terapię logopedyczną w latach 2011-2013.

W przeprowadzonych badaniach oprócz miar statystyki opisowej wykorzystano analizę korespondencji oraz metody wnioskowania statystycznego, takie jak: test serii losowości próby, test normalności Shapiro-Wilka.

2. Materiał i wyniki badań

Pomiar jakości życia przeprowadzono wśród pacjentów jękających się w Centrum Terapii Jękania (CTJ)¹ w latach 2011-2013. Narzędziem badawczym był kwestionariusz jakości życia SF-36v2™ Health Survey, licencja Nr QM009973/2011. Kwestionariusz ten jest narzędziem wystandaryzowanym i ma właściwości psychometryczne.

Kwestionariusz SF-36v2 umożliwia wielowymiarową charakterystykę stanu zdrowia, zawiera 36 pytań (pozycji), diagnozuje 8 aspektów zdrowia. Zawarte

¹ Centrum Terapii Jękania jest prowadzone przez logopedę dypl. Bernadetę Dziekan-Standowicz w Szamocinie przy ul. Parkowej 3 – uczennicę prof. L. Arutiunian (Лилия Зиновьевна Арутюнян) z Moskwy, twórczyni metody „trwałej normalizacji mowy jękających się”. Badanie zostało przeprowadzone za zgodą logopedy dypl. Bernadety Dziekan-Standowicz.

w kwestionariuszu pytania dają ośmiopodskalowy profil, obejmujący zagadnienia zdrowia w zakresie funkcjonalności i dobrostanu [Cooper i in. 2001].

Skala SF-36 składa się z 3 poziomów:

poziom 1 – 36 pozycji;

poziom 2 – 8 podskal, na które składa się od 2 do 10 pozycji;

poziom 3 – 2 wskaźniki sumaryczne, na które składają się poszczególne podskale² [Żołnierczyk-Zreda i in. 2009].

Do badania wylosowano w sposób arbitralny 31 osób spośród 59 pacjentów Centrum. W doborze do próby zastosowano metodę bez zwracania³. Do sprawdzenia założenia dotyczącego losowego charakteru próby posłużono się testem serii losowości próby, w którym sprawdzano hipotezę zerową, że próba ma charakter losowy, wobec hipotezy alternatywnej, która mówi, iż próba nie ma charakteru losowego. W analizie sprawdzono kolejno: losowość (test serii), normalność (test Shapiro-Wilka) [Domański 1990].

Dla tak wylosowanej próby na poziomie istotności $\alpha = 0,05$ nie było podstaw do odrzucenia hipotezy zerowej, że próba ma charakter losowy ($p > 0,05$), zatem można było przyjąć założenie o losowości próby.

Na wstępie dokonano prezentacji tabelarycznej i graficznej analizowanych wyników podskal zmiennych w tab. 1 i rys. 1.

Tabela 1. Struktura badanej próby ze względu na płeć oraz grupy wiekowe (liczba osób)

Wyszczególnienie	Płeć		Razem
	kobieta	mężczyzna	
gw 1 – do 18 lat	2	6	8
gw 2 – od 18 do 25 lat	3	7	10
gw 3 – od 25 do 35 lat	1	7	8
gw 4 – powyżej 35 lat	1	4	5
razem	7	24	31

gw – grupa wiekowa.

Źródło: opracowanie własne.

Podstawowe charakterystyki opisowe zaprezentowano w tab. 2.

Następnie podjęto próbę ustalenia normalności rozkładu wyników dla analizowanych podskal (wyniki na skali 0-100). W celu sprawdzenia hipotezy zerowej o normalności rozkładu wyników analizowanych podskal zastosowano test Shapi-

² Wyniki surowe skali poddawane są operacji rekodowania, a następnie transformacji wyników na skalę 0-100, kolejno dokonuje się transformacji wyników ze skali 0-100 na wyniki znormalizowane o średniej równej 50 i odchyleniu standardowym równym 10. Zaletą wyników znormalizowanych jest możliwość bardzo łatwej interpretacji wyników, ponieważ wszystkie wyniki większe niż 50 są wyższe niż norma, natomiast wyniki niższe niż 50 plasują się poniżej normy. Wynik 100 oznacza najlepsze zdrowie, a wynik 0 najgorsze zdrowie.

³ Wykorzystano generator liczb pseudolosowych w programie Excel.

Tabela 2. Charakterystyki opisowe analizowanych wyników podskal (w pkt)

Wyszczególnienie	Podskala							
	F	R	P	H	V	S	E	W
Minimum	55,00	25,00	40,00	47,00	37,50	12,50	8,33	30,00
Maksimum	100,00	100,00	100,00	100,00	81,25	100,00	100,00	95,00
Kwartył pierwszy	95,00	37,50	62,00	63,50	46,88	50,00	25,00	55,00
Mediana	100,00	75,00	80,00	72,00	62,50	75,00	66,67	60,00
Kwartył trzeci	100,00	93,75	100,00	87,00	68,75	87,50	91,67	70,00
Rozstęp	45,00	75,00	60,00	53,00	43,75	87,50	91,67	65,00
Średnia arytmetyczna	96,51	68,10	78,39	73,94	59,69	67,34	62,68	60,78
Współczynnik skośności	-4,23	-0,52	-0,42	0,05	-0,28	-0,61	-0,27	0,01
Odchylenie standardowe	8,31	28,56	19,80	14,77	13,59	25,34	30,07	14,15
Współczynnik zmienności odchylenia standardowego	8,75%	42,63%	25,68%	20,31%	23,14%	38,25%	48,76%	23,67%
Odchylenie ćwiartkowe	2,5	28,125	19	11,75	10,9375	18,75	33,33333	7,5
Współczynnik zmienności odchylenia ćwiartkowego	2,50%	37,50%	23,75%	16,32%	17,50%	25,00%	50,00%	12,50%

Źródło: opracowanie własne.

ro-Wilka. Sprawdzono, że badana cecha populacji ma rozkład normalny wobec hipotezy alternatywnej, że cecha populacji nie ma rozkładu normalnego. Na rys. 1 przedstawiono histogramy wyników podskal.

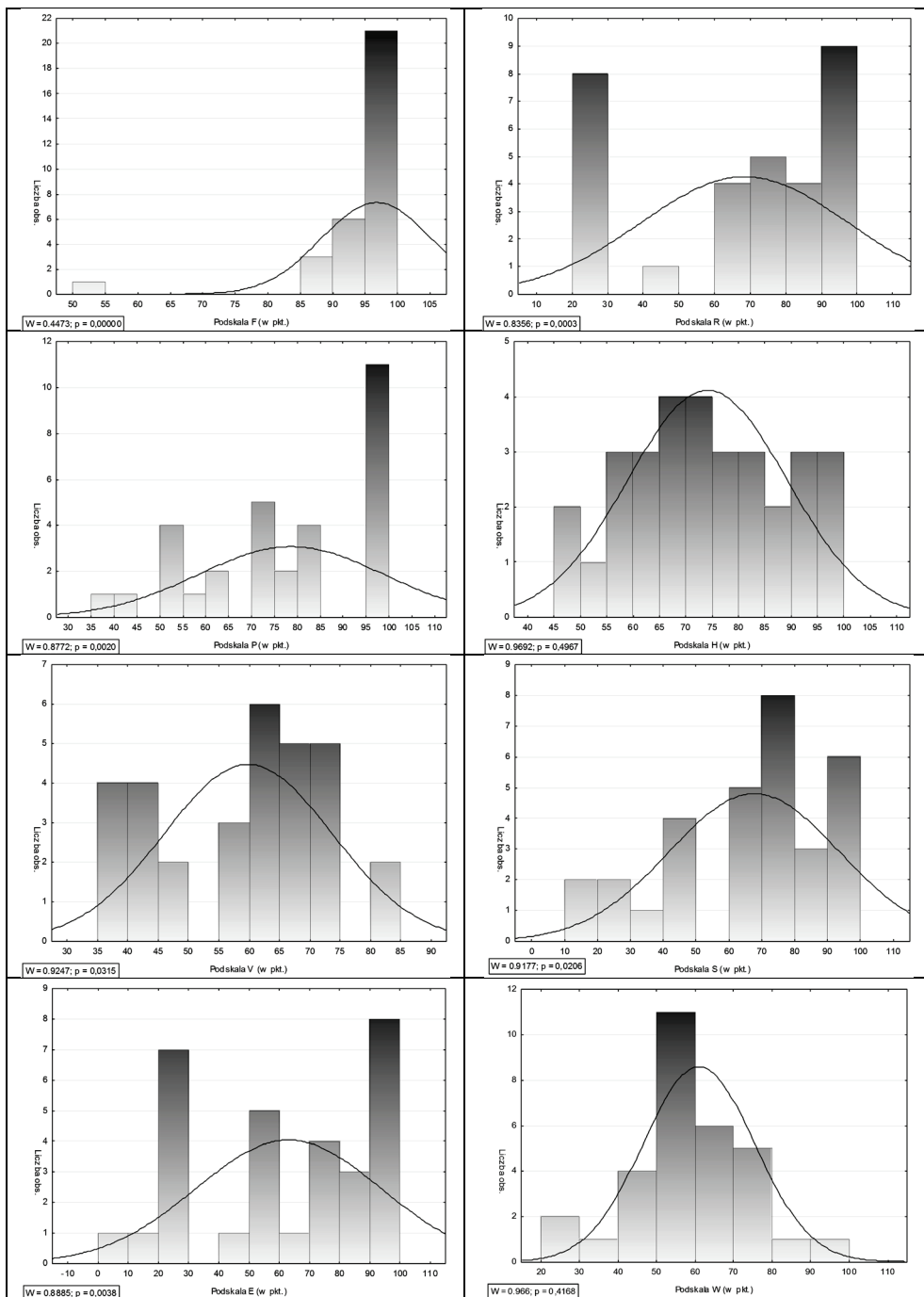
W większości analizowane wyniki podskal charakteryzowały się znaczą asymetrią oraz zróżnicowaniem. Ponadto sprawdzając normalność ich rozkładu testem Shapiro-Wilka, odrzucono hipotezę zerową mówiącą, że rozkład wyników podskali jest normalny również w większości badanych podskal. Tylko dwie spośród wszystkich podskal, spełniały założenie ww. testu, są to podskale: ogólna percepcja zdrowia (H) oraz funkcjonowanie społeczne (S). Rozkład pozostałych wyników podskal nie był normalny. Wyniki dla podskal zestawiono w tab. 3.

Zatem na potrzeby analizy korespondencji utworzono cechy dychotomiczne z analizowanych wyników podskal o kategoriach odpowiednio:

- przekracza medianę w podskali (+);
- nie przekracza mediany w podskali (-).

Badanych pacjentów rozpatrywano ze względu na wiek w następujących grupach:

- 1 – do 18 lat,
- 2 – od 18 do 25 lat,
- 3 – od 25 do 35 lat,
- 4 – powyżej 35 lat.



Rys. 1. Histogramy wyników podskal

Źródło: opracowanie własne.

Tabela 3. Wartości statystyk empirycznych testu Shapiro-Wilka oraz poziomy p

Wyszczególnienie	Podskala F	Podskala R	Podskala P	Podskala H	Podskala V	Podskala S	Podskala E	Podskala W
Statystyka Shapiro-Wilka	0,447	0,836	0,887	0,969	0,925	0,918	0,889	0,966
Poziomy p	0,001	0,001	0,002	0,497	0,032	0,021	0,004	0,417

Źródło: opracowanie własne.

W kolejnym etapie przystąpiono do analizy korespondencji. Analizowane dane zapisano w macierzy Burta (tab. 4).

Tabela 4. Macierz Burta dla grup wiekowych i wyników podskal

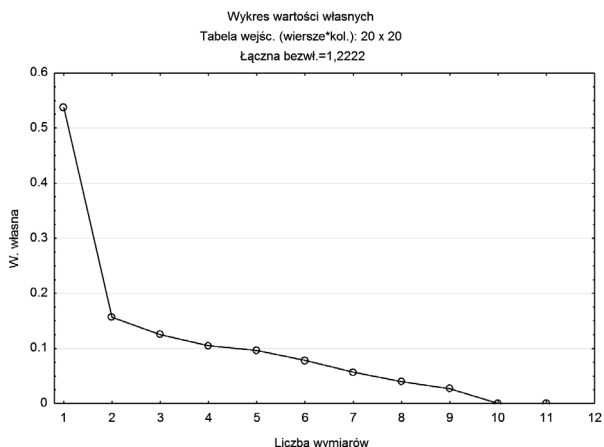
	Tabela zaobs. (Częstości)																				Razem
	Tabela wejśc. (wiersze*kol.): 20 x 20 (Tabl.Burta)																				
	gw 1	gw 2	gw 3	gw 4	R +	R -	P +	P -	H +	H -	V +	V -	S +	S -	E +	E -	W +	W -	F +	F -	
gw:1	8	0	0	0	8	0	8	0	8	0	8	0	8	0	8	0	8	0	6	2	72
gw:2	0	10	0	0	8	2	9	1	8	2	8	2	9	1	8	2	9	1	6	4	90
gw:3	0	0	8	0	7	1	7	1	7	1	8	0	8	0	7	1	8	0	5	3	72
gw:4	0	0	0	5	3	2	5	0	4	1	5	0	4	1	3	2	4	1	4	1	45
R:+	8	8	7	3	26	0	26	0	26	0	26	0	26	0	26	0	26	0	18	8	234
R:-	0	2	1	2	0	5	3	2	1	4	3	2	3	2	0	5	3	2	3	2	45
P:+	8	9	7	5	26	3	29	0	27	2	28	1	28	1	26	3	28	1	20	9	261
P:-	0	1	1	0	0	2	0	2	0	2	1	1	1	1	0	2	1	1	1	1	18
H:+	8	8	7	4	26	1	27	0	27	0	27	0	26	1	26	1	27	0	19	8	243
H:-	0	2	1	1	0	4	2	2	0	4	2	2	3	1	0	4	2	2	2	2	36
V:+	8	8	8	5	26	3	28	1	27	2	29	0	28	1	26	3	28	1	21	8	261
V:-	0	2	0	0	0	2	1	1	0	2	0	2	1	1	0	2	1	1	0	2	18
S:+	8	9	8	4	26	3	28	1	26	3	28	1	29	0	26	3	28	1	20	9	261
S:-	0	1	0	1	0	2	1	1	1	1	1	1	0	2	0	2	1	1	1	1	18
E:+	8	8	7	3	26	0	26	0	26	0	26	0	26	0	26	0	26	0	18	8	234
E:-	0	2	1	2	0	5	3	2	1	4	3	2	3	2	0	5	3	2	3	2	45
W:+	8	9	8	4	26	3	28	1	27	2	28	1	28	1	26	3	29	0	20	9	261
W:-	0	1	0	1	0	2	1	1	0	2	1	1	1	1	0	2	0	2	1	1	18
F:+	6	6	5	4	18	3	20	1	19	2	21	0	20	1	18	3	20	1	21	0	189
F:-	2	4	3	1	8	2	9	1	8	2	8	2	9	1	8	2	9	1	0	10	90
Razem	72	90	72	45	234	45	261	18	243	36	261	18	261	18	234	45	261	18	189	90	2511

Źródło: opracowanie własne.

Następnie, w celu ustalenia, ile wymiarów powinna mieć poszukiwana przestrzeń, posłużono się kryterium osypiska. Znalaziono punkt, gdzie spadek wartości własnych przechodzi do poziomu. Na rys. 2 widać, że osypisko zaczyna się od drugiego wymiaru, zatem wykres sugeruje przyjęcie do analizy przestrzeni dwuwymiarowej.

Wyniki analizy korespondencji wraz z wartościami osobliwymi, wartościami własnymi zestawiono w tab. 5.

Na podstawie powyższych wyników można zauważyć, że kolejne wymiary (które są ortogonalne do pozostałych) wyjaśniają coraz mniejsze części ogólnej



Rys. 2. Wykres osypiska

Źródło: opracowanie własne.

Tabela 5. Wartości osobliwe i wartości własne oraz stopień wyjaśnienia całkowitej bezwładności

Łączna bezwładność = 1,2222					
Liczba wymiarów	Wartości osobliwe	Wartości własne	Procent bezwładności	Procent skumulowany	Chi kwadrat
1	0,7330	0,5373	43,9587	43,9587	398,2692
2	0,3956	0,1565	12,8064	56,7651	116,0266
3	0,3538	0,1252	10,2422	67,0073	92,7951
4	0,3242	0,1051	8,6019	75,6092	77,9335
5	0,3102	0,0962	7,8750	83,4841	71,3477
6	0,2796	0,0782	6,3962	89,8803	57,9498
7	0,2387	0,0570	4,6626	94,5429	42,2433
8	0,1994	0,0398	3,2526	97,7955	29,4692
9	0,1641	0,0269	2,2045	100,0000	19,9725

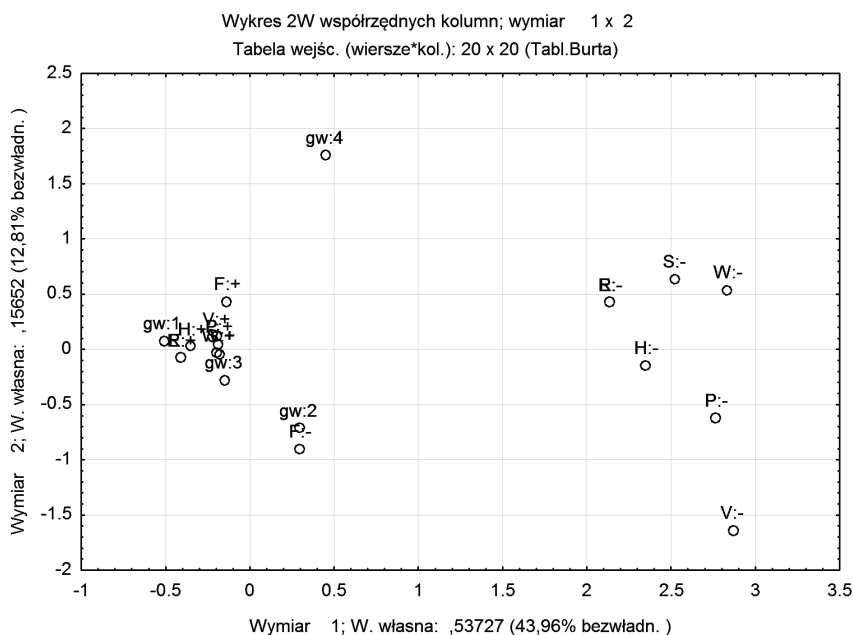
Źródło: opracowanie własne.

wartości statystyki chi-kwadrat (a więc bezwładności). Pierwszy wymiar pozwala na odtworzenie 43,96% całkowitej bezwładności, a uwzględniony drugi wymiar zwiększa procent wyjaśnionej bezwładności do 56,77% całkowitej bezwładności. Zatem na podstawie tego kryterium wskazane jest usytuowanie profili w przestrzeni dwuwymiarowej.

Po ustaleniu liczby wymiarów w następnym kroku obliczono współrzędne profili kolumnowych w nowym, ortonormalnym układzie współrzędnych, wyznaczonym przez wektory osobliwe. Do interpretacji współrzędnych punktów reprezentujących kolumny zastosowano metodę standaryzacji wierszowo-kolumnowej, gdzie współrzędne wierszy obliczane są z macierzy profili wierszy, a współrzędne kolumn z macierzy profili kolumn. Standaryzacja ta pozwoliła na uzyskanie jedno-

czeńnie współrzędnych punktów reprezentujących grupy wiekowe oraz wyniki podskal.

Oś pozioma ma największy (43,96%) udział w bezwładności. Należy zatem do najważniejszego wymiaru wyjaśniającego większość zróżnicowania pomiędzy kolumnami. Na wykresie ta oś wyróżnia dwie grupy. Najbardziej na lewo względem środka tej osi znajdują się osoby w wieku do 18 lat (gw 1) oraz od 25 do 35 lat (gw 3), których wyniki w ośmiu analizowanych podskalach są powyżej mediany. Natomiast na prawo względem środka tej osi są osoby w wieku od 18 do 25 lat (gw 2) oraz powyżej 35 lat (gw 4), posiadające wyniki w ośmiu analizowanych podskalach poniżej mediany.



Rys. 3. Powiązania pomiędzy kategoriami zmiennych

Źródło: opracowanie własne.

Druga oś pionowa dzieli badanych na trzy grupy. Najbardziej u dołu od środka tej osi znajdują się osoby w wieku od 18 do 25 lat (gw 2), dla których wynik uzyskany w podskali sprawność fizyczna (F) jest poniżej mediany. Grupy druga i trzecia znajdują się blisko środka drugiej osi, wśród nich są osoby do 18 lat (gw 1), których wyniki, będące powyżej mediany, są uzyskane w podskalach: ograniczenia aktywności z powodu zdrowia fizycznego (R+), ogólna percepcja zdrowia (H+), ograniczenie aktywności z powodu problemów emocjonalnych (E+). Do trzeciej grupy należą osoby od 25 do 35 lat (gw 3). Ich wyniki są powyżej mediany w podskalach: witalność (V+), dolegliwości bólowe (P+), zdrowie psychiczne (W+),

funkcjonowanie społeczne (S+), sprawność fizyczna (F+). Usytuowanie grupy drugiej i trzeciej blisko środka osi oznacza, że profil dla tych kolumn jest zbliżony do profilu przeciętnego.

Tabela 6. Współrzędne kolumn i statystyki jakości rozwiązania

Kolumna nazwa	Kolumna Liczba	Współrz. Wymiar1	Współrz. Wymiar2	Masa	Jakość	Względna bezwład.	Bezwład. Wymiar1	Cos ² - Wymiar1	Bezwład. Wymiar2	Cos ² Wymiar2
gw:1	1	-0,5108	0,0779	0,0287	0,0928	0,0674	0,0139	0,0907	0,0011	0,0021
gw:2	2	0,2987	-0,7179	0,0358	0,2879	0,0616	0,0060	0,0425	0,1180	0,2454
gw:3	3	-0,1463	-0,2793	0,0287	0,0346	0,0674	0,0011	0,0074	0,0143	0,0271
gw:4	4	0,4540	1,7581	0,0179	0,6341	0,0762	0,0069	0,0396	0,3539	0,5944
F:+	5	-0,1404	0,4307	0,0753	0,4309	0,0293	0,0028	0,0414	0,0892	0,3895
F:-	6	0,2948	-0,9044	0,0358	0,4309	0,0616	0,0058	0,0414	0,1873	0,3895
R:+	7	-0,4107	-0,0813	0,0932	0,9116	0,0147	0,0293	0,8773	0,0039	0,0343
R:-	8	2,1359	0,4225	0,0179	0,9116	0,0762	0,1522	0,8773	0,0204	0,0343
P:+	9	-0,1907	0,0432	0,1039	0,5542	0,0059	0,0070	0,5271	0,0012	0,0271
P:-	10	2,7645	-0,6270	0,0072	0,5542	0,0850	0,1020	0,5271	0,0180	0,0271
H:+	11	-0,3482	0,0219	0,0968	0,8219	0,0117	0,0218	0,8186	0,0003	0,0032
H:-	12	2,3507	-0,1481	0,0143	0,8219	0,0792	0,1474	0,8186	0,0020	0,0032
V:+	13	-0,1977	0,1132	0,1039	0,7526	0,0059	0,0076	0,5668	0,0085	0,1858
V:-	14	2,8668	-1,6415	0,0072	0,7526	0,0850	0,1097	0,5668	0,1234	0,1858
S:+	15	-0,1738	-0,0438	0,1039	0,4660	0,0059	0,0058	0,4381	0,0013	0,0279
S:-	16	2,5205	0,6355	0,0072	0,4660	0,0850	0,0848	0,4381	0,0185	0,0279
E:+	17	-0,4107	-0,0813	0,0932	0,9116	0,0147	0,0293	0,8773	0,0039	0,0343
E:-	18	2,1359	0,4225	0,0179	0,9116	0,0762	0,1522	0,8773	0,0204	0,0343
W:+	19	-0,1955	-0,0371	0,1039	0,5740	0,0059	0,0074	0,5540	0,0009	0,0200
W:-	20	2,8343	0,5383	0,0072	0,5740	0,0850	0,1072	0,5540	0,0133	0,0200

Źródło: opracowanie własne.

Wśród wszystkich kolumn najbardziej ważne okazały się kolumny reprezentujące wyniki podskal: dolegliwości bólowe (P+), witalność (V+), funkcjonowanie społeczne (S+), zdrowie psychiczne (W+), ponieważ masa tych kolumn jest największa.

Najlepiej odwzorowane w przestrzeni dwuwymiarowej są wyniki dla podskal: ograniczenia aktywności z powodu zdrowia fizycznego (R+, R-), ograniczenie aktywności z powodu problemów emocjonalnych (E+, E-), ogólna percepcja zdrowia (H+), ponieważ jakość tych kolumn jest bliska jedności.

Wyniki takich podskal, jak: dolegliwości bólowe (P-), witalność (V-), funkcjonowanie społeczne (S-), zdrowie psychiczne (W-), ogólna percepcja zdrowia (H-), mają największy udział w ogólnej bezwładności w pierwotnej wielowymiarowej przestrzeni wśród kolumn.

Do zdefiniowania pierwszego wymiaru w największym stopniu przyczyniły się wyniki podskal: ograniczenia aktywności z powodu zdrowia fizycznego (R-), ograniczenie aktywności z powodu problemów emocjonalnych (E-), ogólna percepcja zdrowia (H-), o czym świadczą najwyższe absolutne udziały tych kolumn w bezwładności pierwszego wymiaru.

Pierwszy wymiar w największym stopniu jest opisywany przez punkty reprezentujące podskale: ograniczenia aktywności z powodu zdrowia fizycznego (R+, R-), ograniczenie aktywności z powodu problemów emocjonalnych (E+, E-), o czym świadczy wysoka wartość absolutnego udziału pierwszego wymiaru w bezwładności tych podskal.

Do zdefiniowania drugiego wymiaru w największym stopniu przyczyniły się osoby powyżej 35 lat (gw 4) oraz podskala sprawność fizyczna (F+), o czym świadczą najwyższe absolutne udziały tych kolumn w bezwładności drugiego wymiaru.

Drugi wymiar w największym stopniu jest opisywany przez punkty reprezentujące podskalę sprawność fizyczna (F+, F-) oraz osoby powyżej 35 lat (gw 4), o czym świadczą wysokie wartości absolutnych udziałów drugiego wymiaru w bezwładności.

3. Podsumowanie

Wyniki przeprowadzonej analizy wskazują na możliwość zastosowania analizy korespondencji do oceny jakości życia wśród jaskających się pacjentów w takich podskalach, jak: sprawność fizyczna, ograniczenia aktywności z powodu zdrowia fizycznego, dolegliwości bólowe, ogólna percepcja zdrowia, witalność, funkcjonowanie społeczne, zdrowie psychiczne oraz ograniczenie aktywności z powodu problemów emocjonalnych. Poza tym dostarczyły nam informacji o strukturze powiązań między badanymi podskalami jakości życia a grupami wiekowymi.

Na podstawie przeprowadzonego badania okazało się, że oryginalne konfiguracje punktów reprezentujących podskale i grupy wiekowe zostały przedstawione w dwuwymiarowej przestrzeni. Spośród grup wiekowych najlepiej odwzorowane zostały w przestrzeni dwuwymiarowej grupy do 18 lat i od 18 do 25 lat, a w przypadku podskal jakości życia: ograniczenie aktywności z powodu zdrowia fizycznego (R), ogólna percepcja zdrowia (H) oraz funkcjonowanie społeczne (S).

Najlepiej odwzorowane w przestrzeni dwuwymiarowej są wyniki dla podskal: ograniczenia aktywności z powodu zdrowia fizycznego (R+, R-), ograniczenie aktywności z powodu problemów emocjonalnych (E+, E-), ogólna percepcja zdrowia (H+).

Literatura

- Cooper J.K., Kohlmann T., Michael J.A. i in. (2001), *Health outcomes. New quality measure for Medicare*, „International Journal for Quality in Health Care”, 13(1), s. 13-16.
- Domański C. (1990), *Testy statystyczne*, PWE, Warszawa.
- Greenacre M. (1984), *Theory and applications of correspondence analysis*, Academic Press, London.
- Stanimir A. (2005), *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Tarkowski Z. (2001), *Jąkanie*, PWN, Warszawa.

Walesiak M., Gatnar E. (red.) (2009), *Statystyczna analiza danych z wykorzystaniem programu R*, PWN, Warszawa.

WHOQOL Group (1993), *Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL)*, *Quality Life Research*, 2, s.153-159.

Żołnierczyk-Zreda D., Wrzeźniewski K., Bugajska J., Jędryka-Góral A. (2009), *Polska wersja kwestionariusza SF-36v2 do badania jakości życia*, CIOP-PIB, Warszawa.

APPLICATION OF CORRESPONDENCE ANALYSIS FOR THE MEASUREMENT OF QUALITY OF LIFE – QUESTIONNAIRE SF-36V2 BASED RESEARCH

Summary: The aim of this article is an application of correspondence analysis to assess the quality of life of people interviewed with the quality of life questionnaire SF-36v2™ Health Survey in eight subscales: physical fitness, activity limitations due to physical health, pain, general health perception, vitality, social functioning, mental health and activity limitation due to emotional problems. The data was collected for the period 2011-2013.

Keywords: quality of life, correspondence analysis, the quality of life questionnaire SF-36v2™.