

# 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](http://www.ibuk.pl), [www.ebscohost.com](http://www.ebscohost.com),

w Dolnośląskiej Bibliotece Cyfrowej [www.dbc.wroc.pl](http://www.dbc.wroc.pl),

The Central and Eastern European Online Library [www.ceeol.com](http://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](http://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

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

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

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

## Summaries

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

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

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

**Joanna Banaś, Małgorzata Machowska-Szewczyk**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

**Bożena Mroczek**

Pomorski Uniwersytet Medyczny w Szczecinie

---

## ZASTOSOWANIE ANALIZY KORESPONDENCJI DO BADANIA WPŁYWU ELEKTROWNI WIATROWYCH NA JAKOŚĆ ŻYCIA LUDNOŚCI

---

**Streszczenie:** W artykule przedstawiono wyniki poszukiwania zależności między jakością życia mieszkańców miejscowości położonych w strefie oddziaływania farm wiatrowych na terenie Polski a obecnością farm wiatrowych na różnym etapie inwestycji. Do badania wpływu obecności farm wiatrowych na jakość życia zastosowano analizę korespondencji, która umożliwia analizę nie tylko danych ilościowych, ale także mierzonych na skalach nominalnych, porządkowych oraz nie stawia wymagań co do liczebności zbioru jednostek statystycznych. Metoda ta jest stosowana w celu wydobywania wiedzy ze zbiorów danych poprzez analizę częstości współwystępowania poszczególnych wariantów obserwowanych cech.

**Słowa kluczowe:** analiza korespondencji, kwestionariusz SF-36v2, elektrownie wiatrowe.

### 1. Wstęp

Elektrownie wiatrowe wytwarzają energię elektryczną za pomocą generatorów (turbiny wiatrowych) napędzanych energią wiatru. Uzyskana energia elektryczna jest uznawana za ekologicznie czystą, gdyż jej wytworzenie nie pociąga za sobą spalania żadnego paliwa. Budowa elektrowni wiatrowych zawsze budzi wiele kontrowersji. Przeciwnicy energetyki wiatrowej uważają, że turbiny są źródłem hałasu i przez to ich praca może być uciążliwa dla mieszkańców najbliższych okolic. Z kolei zwolennicy wskazują na elektrownie wiatrowe jako źródło energii odnawialnej, która nie emituje do atmosfery żadnych zanieczyszczeń.

Celem badania jest poszukiwanie zależności między oceną jakości życia przez mieszkańców miejscowości położonych w strefie oddziaływania farm wiatrowych na terenie Polski a obecnością farm wiatrowych na różnym etapie inwestycji. Jakość życia oceniono na podstawie skali SF-36v2, skali odczuwania problemów



zdrowotnych oraz autorskiego kwestionariusza ankiety uwzględniającego pytania o dane socjodemograficzne i występujące choroby przewlekłe [Mroczek 2013].

Do badania wpływu obecności farm wiatrowych na różnym etapie inwestycji na jakość życia zastosowano analizę korespondencji, która jest wyspecjalizowaną metodą eksploracji danych, ułatwiającą prezentację – przede wszystkim w postaci graficznej – związków pomiędzy cechami oraz pomiędzy jednostkami statystycznymi. Umożliwia ona analizę nie tylko danych ilościowych, ale także mierzonych na skalach nominalnych, porządkowych oraz nie stawia wymagań co do liczebności zbioru jednostek statystycznych. Metoda ta jest stosowana w celu wydobywania wiedzy ze zbiorów danych poprzez analizę częstości współwystępowania poszczególnych wariantów obserwowanych cech.

Wyniki badań stanowią interesujące uzupełnienie uzyskanych dotychczas rezultatów.

## 2. Opis metody

Analiza korespondencji jest narzędziem służącym do badania relacji zachodzących pomiędzy obiektami lub kategoriami cech typu jakościowego. Algorytm badania można przedstawić w następujących krokach<sup>1</sup>:

1. Skonstruowanie odpowiedniej macierzy lub tabeli danych. Do zapisu danych najczęściej wykorzystuje się macierz Burta, która jest wynikiem działania:

$$\mathbf{B} = \mathbf{Z}^T \mathbf{Z} = [b_{ij}]_{J \times J}, \quad J = \sum_{q=1}^Q J_q,$$

gdzie:  $J_q$  to liczba kategorii cechy  $X_q$ ,

$\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_Q]$  – złożona macierz znaczników dla  $n$  obiektów i  $Q$  zmiennych.

Elementami podmacierzy  $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_Q$ , odpowiadających poszczególnym cechom, są wartości 1 lub 0, w zależności od tego, czy dany obiekt posiada wyróżnioną kategorię ustalonej cechy, czy też nie.

2. Wyznaczenie macierzy zaobserwowanych częstości względnych  $\mathbf{P} = [p_{ij}]_{J \times J}$ , której elementy otrzymuje się, dzieląc zaobserwowane liczebności w poszczególnych komórkach macierzy  $\mathbf{B}$  przez sumę liczebności wszystkich cel macierzy Burta, czyli  $nQ^2$ .

3. Wyznaczenie profili wierszowych (kolumnowych) jako ilorazów zaobserwowanych częstości względnych wiersza (kolumny) przez sumę częstości danego wiersza (kolumny). Średnie profile wierszowe oraz kolumnowe zwane są, odpo-

<sup>1</sup> Szczegółowy opis metody można znaleźć w pracach: [Stanimir 2005; Sompolska-Rzechuła 2010; Balicki 2009; Panek 2009].

wiednio, centrum wierszowym lub kolumnowym. Centrum wierszowe lub kolumnowe stanowi punkt przecięcia osi przestrzeni rzutowania.

4. Przekształcenie macierzy  $\mathbf{P}$  w macierz  $\mathbf{A}$  ważonych odchyłeń profili od centrum wierszowego (kolumnowego), zwaną macierzą różnic zestandaryzowanych w postaci:

$$\mathbf{A} = [a_{ij}], \text{ gdzie: } a_{ij} = \frac{p_{ij} - p_{i\bullet} p_{\bullet j}}{\sqrt{p_{i\bullet} p_{\bullet j}}},$$

$p_{ij}$  – elementy macierzy  $\mathbf{P}$ ,  $p_{i\bullet} = \frac{Qb_{ii}}{nQ^2}$ ,  $p_{\bullet j} = \frac{Qb_{jj}}{nQ^2}$ ,  $i, j \in \{1, 2, \dots, J\}$ .

5. Dokonanie rozkładu macierzy  $\mathbf{A}$  według wartości osobliwych, czyli:

$$\mathbf{A} = \mathbf{U}\mathbf{\Gamma}^2\mathbf{U}^T,$$

gdzie:  $\mathbf{U}$  – macierz wektorów własnych macierzy  $\mathbf{A}$ ,

$\mathbf{\Gamma}^2 = [\gamma_k^2]$  – macierz diagonalna zawierająca kwadraty wartości osobliwych macierzy  $\mathbf{A}$ ,  $k \in \{1, 2, \dots, K\}$ ,  $K = \sum_{q=1}^Q (J_q - 1)$ ,  $J_q$  – liczba kategorii cechy  $X_q$ .

6. Wyznaczenie macierzy zawierającej współrzędne kategorii cech na wszystkich osiach rzeczywistych powiązań:

$$\mathbf{F} = \mathbf{D}_r \mathbf{U} \mathbf{\Gamma},$$

gdzie  $\mathbf{D}_r$  – macierz diagonalna odwrotności pierwiastków częstości brzegowych wierszy. Wymiar przestrzeni rzeczywistych powiązań wynosi  $K$ .

7. Poszukiwanie niższego wymiaru przestrzeni do prezentacji wyników tak, aby zniekształcenie wyjściowej konfiguracji punktów było jak najmniejsze.

a) Kryterium wyboru wymiaru przestrzeni rzutowania może być udział inercji ustalonego wymiaru w inercji całkowitej, czyli sumie wszystkich inercji głównych:

$$\tau_{K^*} = \frac{\sum_{k=1}^{K^*} \lambda_k}{\sum_{k=1}^K \lambda_k} = \frac{\sum_{k=1}^{K^*} \lambda_k}{\lambda},$$

gdzie:  $K^* \in \{1, \dots, K\}$  oznacza wybrany wymiar rzutowania,  $\lambda_k = \gamma_k^2$ ,  $k \in \{1, \dots, K\}$  są wartościami własnymi przestrzeni rzutowania. Wybiera się wówczas takie  $K^*$ , dla którego stopień wyjaśniania inercji ma wartość bliską 1, a zwiększanie wymiaru przestrzeni nie powoduje gwałtownego wzrostu tego udziału.

b) Greenacre [1993] zaproponował wyznaczenie optymalnego wymiaru  $K^*$  przestrzeni następująco:

$$K^* = \max \left\{ k \in \{1, \dots, K\} : \lambda_k > \frac{1}{Q} \right\}$$

oraz poprawę wyników analizy zmiennych zapisanych w macierzy Burta przez przeskalowanie współrzędnych poszczególnych punktów:

$$\tilde{\mathbf{F}} = \mathbf{F}^* \mathbf{\Gamma}^{-1} \tilde{\mathbf{\Lambda}},$$

gdzie  $\tilde{\mathbf{\Lambda}}$  jest macierzą diagonalną pierwszych  $K^*$  wartości własnych zmodyfikowanych według wzoru:

$$\tilde{\lambda}_k = \left( \frac{Q}{Q-1} \right)^2 \left( \sqrt{\lambda_k} - \frac{1}{Q} \right)^2,$$

$\mathbf{F}^*$  jest macierzą pierwszych  $K^*$  współrzędnych kategorii z macierzy  $\mathbf{F}$ .

c) Kryterium łockia polega na określeniu wymiaru przestrzeni na podstawie wykresu uporządkowanych nierosnąco wszystkich niezerowych wartości własnych. Numer wartości własnej, dla której nastąpiło zagięcie, wskazuje na wymiar przestrzeni rzutowania.

Wyniki analizy korespondencji prezentuje się w przestrzeni maksymalnie trójwymiarowej przy zachowaniu możliwie najmniejszej straty informacji o różnicowaniu częstości występowania poszczególnych wariantów cech i ocenia położenie punktów względem centrum rzutowania lub innych punktów.

Gdy liczba wariantów jest duża i wyodrębnienie klas na podstawie rozrzutu punktów na wykresie jest niejednoznaczne lub najlepszą formą prezentacji współwystępowania cech jest przestrzeń o wymiarze większym niż trzy, wykorzystuje się wybrane metody klasyfikacji. Jako obiekty przyjmuje się wszystkie kategorie analizowanych cech, natomiast jako zmienne – wartości współrzędnych rzutowania każdej kategorii. Do prezentacji wyników analizy korespondencji w pracy zastosowano ocenianą jako najbardziej efektywną metodę Warda [Ward 1963; Gordon 1999; Ostasiewicz (red.) 1998].

### 3. Badanie ankietowe

Do badania jakości życia zastosowano Skalę SF-36v2, która składa się z 36 pytań tak przekształconych, by można było uzyskać osiem domen jakości życia (*QoL*). Odpowiedzi punktowano w skali od 0 do 100, przy czym wyższe średnie wyniki wskazują lepszą jakość życia i lepszy stan zdrowia. Na wykorzystanie Skali SF-36v2 uzyskano licencję Office of Grants Scholary Research. Zastosowano również skalę odczuwania problemów zdrowotnych oraz autorski kwestionariusz ankiety, uwzględniający pytania o dane socjodemograficzne, ryzykowne zachowania zdrowotne i występujące choroby przewlekłe.

Badania przeprowadzono w grupie 1277 mieszkańców województw, w których w roku 2010 zlokalizowanych było najwięcej farm wiatrowych lub rozpoczęto związane z nimi inwestycje. Były to województwa: zachodniopomorskie, pomorskie, wielkopolskie, dolnośląskie, warmińsko-mazurskie, podlaskie i podkarpackie [Mroczek 2013]. Po usunięciu brakujących danych w analizie uwzględniono 1041 obiektów. Samoocenę jakości zdrowia ankietowanych reprezentuje zmienna syntetyczna *QoL* (*quality of life*) powstała jako średnia arytmetyczna 8 domen (funkcjonowanie fizyczne, ból fizyczny, ograniczenia fizyczne w pełnieniu ról, ograniczenia emocjonalne w pełnieniu ról, funkcjonowanie społeczne, ogólne postrzeganie zdrowia, zdrowie psychiczne, energia i witalność), wyznaczonych na podstawie pytań z odpowiedziami w skali Likerta i zakodowanych liczbowo. Wartości średnie zostały skategoryzowane w sposób liniowy i równomierny do wariantów: 0, 25, 50, 75, 100, odpowiadających 5-stopniowej skali Likerta.

Do określenia czynników wpływających na ocenę jakości życia przyjęto następujący zestaw cech i odpowiadających im kategorii:

1. *QoL* – jakość życia (0; 25; 50; 75; 100),
2. *Wiek* – (do 30; 31-40; 41-50; 51-60; od 61),
3. *Płeć* – (K – kobieta; M – mężczyzna),
4. *Wyk* – wykształcenie (P – podstawowe i niżej; ZZ – zasadnicze zawodowe; S – średnie, w tym policealne; W – wyższe);
5. *Akt\_zaw* – aktywność zawodowa (P – pracujący; B – bezrobotny, U – uczeń/student; E – renta/emerytura),
6. *Odl\_F* – odległość miejsca zamieszkania od farmy wiatrowej (pon\_700 – poniżej 700 m; 700-1000; 1001-1500; 1501-2000; niezn – nie potrafię określić, ale nie więcej niż 2 km),
7. *EWi* – obecność energetyki wiatrowej (T – tak – inwestycja w trakcie realizacji, N – nie – inwestycja planowana; W – wiatraki stoją; Nw – nie wiem),
8. *Alk* – picie alkoholu (R – rzadko lub wcale; Cz – często),
9. *Pap* – palenie papierosów (N – nie palę lub okazjonalnie; T – palę regularnie),
10. *Doleg\_B* – częstotliwość odczuwania dolegliwości: bólu głowy, brzucha lub pleców, w ciągu ostatnich 6 miesięcy (R – rzadko; Cz – często),
11. *Doleg\_Roz* – częstotliwość odczuwania przygnębienia, rozdrażnienia, złości, zdenerwowania w ciągu ostatnich 6 miesięcy (R – rzadko; Cz – często),
12. *Doleg\_Zas* – częstotliwość odczuwania trudności w zasypianiu, lęku, niepokoju w ciągu ostatnich 6 miesięcy (R – rzadko; Cz – często),
13. *Doleg\_Zm* – częstotliwość odczuwania silnego zmęczenia, wyczerpania w ciągu ostatnich 6 miesięcy (R – rzadko; Cz – często),
14. *Chor* – leczenie z powodu choroby przewlekłej: cukrzyca, nadciśnienia tętnicze, choroba wieńcowa, choroba reumatyczna, choroby płuc, choroba nowotworowa, choroba alkoholowa, inna niewymieniona (T – tak; N – nie).

## 4. Wyniki badania

Do oceny współwystępowania kategorii cechy jakość życia oraz pozostałych cech przyjętych w badaniu za pomocą metody wielowymiarowej analizy korespondencji wykorzystano program komputerowy MS Excel 2007 oraz pakiet Statistica 8.0 PL.

Przed zastosowaniem wielowymiarowej analizy zgodności zbadano testem niezależności  $\chi^2$ , czy istnieje statystycznie istotna zależność pomiędzy kategoriami cechy zależnej a pozostałymi cechami. Z dalszej analizy wyeliminowano cechy: *Pleć* oraz *Doleg\_Roz*, które nie są powiązane z cechą *QoL*. W przypadku badania zależności pomiędzy kategoriami cechy zależnej (*QoL*) a kategoriami cech z nią powiązanych, otrzymano macierz Burtę o wymiarach  $39 \times 39$ . Wymiar rzeczywistej przestrzeni współwystępowania kategorii cech wynosił 27. Następnie poszukiwano niższego wymiaru przestrzeni wspólnej dla profili wierszowych oraz kolumnowych, tak aby możliwie duża część inercji całkowitej była wyjaśniona. Dokonano rozkładu macierzy różnic zestandaryzowanych według wartości osobliwych i zbadano, w jakim stopniu całkowita bezwładność została wyjaśniona przez wartości własne przestrzeni o niższym wymiarze. W tabeli 1 przedstawiono część wartości własnych  $\lambda_k$  macierzy zestandaryzowanych różnic (kwadraty wartości osobliwych  $\gamma_k$ ), procent bezwładności  $\lambda_k/\lambda$  oraz procent skumulowany, tj. udział wartości własnych z wymiaru  $K$  w inercji całkowitej.

**Tabela 1.** Wartości osobliwe i wartości własne oraz stopień wyjaśnienia całkowitej bezwładności

Liczba wymiarów	Wartości osobliwe $\gamma_k$	Wartości własne $\lambda_k$	Procent bezwładności $\lambda_k/\lambda$	Skumulowany procent
1	0,5016	0,2516	11,1806	11,1806
2	0,3733	0,1393	6,1931	17,3738
3	0,3542	0,1254	5,5743	22,9481
...	...	...	...	...
29	0,1767	0,0312	1,3869	100,0000
		$\lambda = 2,250$		

Źródło: opracowanie własne.

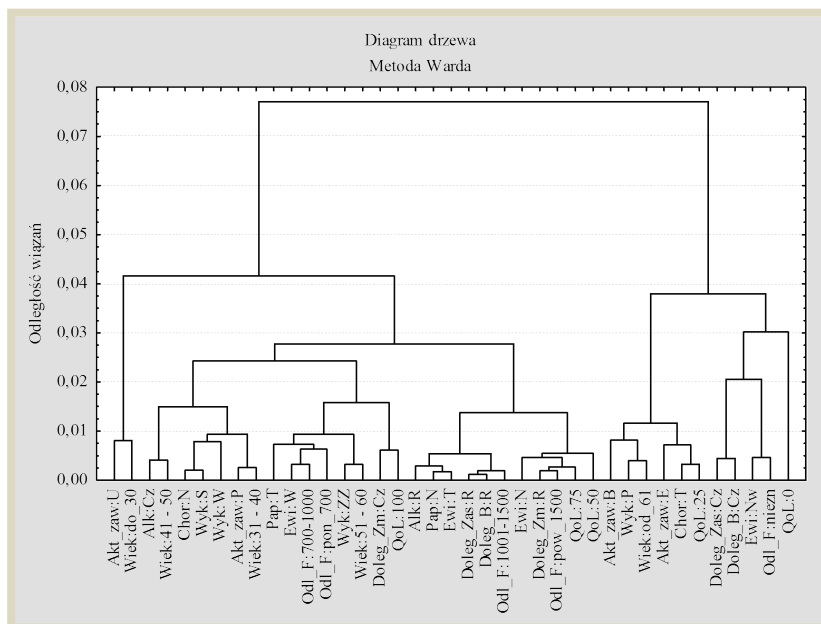
Posługując się kryterium Greenacre'a za najlepszą przestrzeń rzutowania należy uznać  $\mathbb{R}^{10}$ . W tabeli 2 przedstawiono zmodyfikowane zgodnie z propozycją Greenacre'a wartości osobliwe oraz własne. Analizując zmodyfikowane wartości własne oraz udział w bezwładności całkowitej inercji poszczególnych wymiarów (tab. 2), do graficznej prezentacji współwystępowania kategorii cech wybrano przestrzeń czterowymiarową, która odwzorowuje około 52% łącznej bezwładności.

Do wskazania powiązań między kategoriami poszczególnych cech wybrano metodę Warda, za pomocą której utworzono dendrogram przedstawiony na rys. 1.

**Tabela 2.** Zmodyfikowane wartości osobliwe i wartości własne oraz stopień wyjaśnienia całkowitej bezwładności

Liczba wymiarów	Wartości osobliwe $\gamma_k$	Wartości własne $\lambda_k$	Procent bezwładności $\lambda_k/\lambda$	Skumulowany procent
1	0,0456	0,0021	20,4449	20,4449
2	0,0339	0,0012	11,3248	31,7696
3	0,0322	0,0010	10,1932	41,9628
4	0,0319	0,0010	10,0349	51,9978
5	0,0310	0,0010	9,4288	61,4266
6	0,0289	0,0008	8,1948	69,6214
7	0,0285	0,0008	7,9802	77,6017
8	0,0284	0,0008	7,9417	85,5434
9	0,0274	0,0008	7,3744	92,9178
10	0,0268	0,0007	7,0822	100,0000
		$\lambda = 0,0102$		

Źródło: opracowanie własne.

**Rys. 1.** Diagram podziału kategorii cech za pomocą metody Warda

Źródło: opracowanie własne.

Zaproponowano podział na cztery klasy, wyróżniając dwa skupienia w klasie drugiej. Pierwsza klasa wskazuje na dość oczywiste powiązanie bycia studentem z wiekiem do 30 lat.

W drugiej klasie można w celu wygodniejszej interpretacji wyróżnić dwa skupienia. Najwyższy poziom oceny jakości życia jest związany z brakiem występowania poważnych chorób, ale również częstym odczuwaniem zmęczenia, pracą zawodową w wieku powyżej 30 lat i wykształceniem powyżej podstawowego, paleniem papierosów i częstym piciem alkoholu, obecnością farm wiatrowych w odległości do 1 km od miejsca zamieszkania. Średni lub wyższy poziom oceny jakości życia jest powiązany z rzadkim występowaniem wszystkich branych pod uwagę w kwestionariuszu problemów zdrowotnych, niepaleniem papierosów i unikaniem alkoholu, występowaniem farm wiatrowych w odległości powyżej 1 km od miejsca zamieszkania.

Niski poziom oceny jakości życia (klasa trzecia) jest związany z podeszłym wiekiem (powyżej 60 lat), brakiem aktywności zawodowej (brak pracy lub renta/emerytura), wykształceniem podstawowym i występowaniem poważnych chorób.

Najniższy poziom oceny jakości życia współwystępuje z częstym pojawianiem się bólu głowy, brzucha lub pleców w ciągu ostatnich 6 miesięcy, częstym odczuwaniem trudności w zasypianiu, lęku, niepokoju w ciągu ostatnich 6 miesięcy oraz brakiem wiedzy na temat farm wiatrowych w pobliżu miejsca zamieszkania.

## 5. Podsumowanie

Celem badania było ustalenie, czy obecność farm wiatrowych ma wpływ na ocenę jakości życia osób mieszkających w strefie oddziaływania farm wiatrowych. Zastosowana analiza korespondencji pozwala stwierdzić, że obecność elektrowni wiatrowych w pobliżu miejsca zamieszkania nie wpływa negatywnie na ocenę tej jakości. Więcej, najwyższe oceny standardu życia współwystępują z obecnością farm wiatrowych na różnym etapie inwestycji w odległości nie większej niż 1 km od miejsca zamieszkania, a więc w najbliższym sąsiedztwie. Mimo występujących wśród mieszkańców obaw związanych z obecnością turbin wiatrowych, a dotyczących negatywnego wpływu na zdrowie człowieka i stan środowiska, nie można ich na podstawie tego badania uzasadnić. Wydaje się, że wskazane byłoby powtórzenie badań za kilkanaście lat i porównanie ocen jakości życia w odniesieniu do obecności farm wiatrowych.

## Literatura

- Balicki A. (2009), *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
- Gordon A.D. (1999), *Classification*, Chapman & Hall/CRC, Boca Raton.
- Greenacre M. (1993), *Correspondence Analysis in Practice*, London, Academic Press.
- Mroczek B. (2013), *Jakość życia osób mieszkających w strefie oddziaływania farm wiatrowych*, Rozprawa habilitacyjna wykonana z Zakładu Zdrowia Publicznego Wydział Nauk o Zdrowiu PUM w Szczecinie, Wydawnictwo ZAPOL, Szczecin.

- Ostasiewicz W. (red.) (1998), *Statystyczne metody analizy danych*, Akademia Ekonomiczna we Wrocławiu, Wrocław.
- Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, Szkoła Główna Handlowa w Warszawie, Warszawa.
- Stanimir A. (2005), *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- Sompolska-Rzechuła A. (2010), *Zastosowanie metody analizy zgodności w badaniu jakości życia kobiet*, „Wiadomości Statystyczne” nr 1, Warszawa.
- Ward J.H. (1963), *Hierarchical grouping to optimize an objective function*, „Journal of the American Statistical Association”, No. 58.

## USING THE CORRESPONDENCE ANALYSIS TO EXAMINE THE IMPACT OF WIND TURBINES ON THE QUALITY OF LIFE

**Summary:** In the article the results of seeking the relation between the quality of life of the residents and the presence of wind farms are shown. To study the impact of the wind farms' presence on the quality of life a correspondence analysis was used, which allows to analyze not only quantitative data but also the ones measured on nominal or ordinal scales, and does not set requirements for the number of statistical units. This method is applied in order to take the knowledge from the data sets through the frequency analysis of the co-occurrence of observed features' individual variants.

**Keywords:** correspondence analysis, SF-36v2 questionnaire, wind turbines.