

# PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

# RESEARCH PAPERS

of Wrocław University of Economics

Nr 327

**Taksonomia 22**

**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> .....  | 9   |
| <b>Eugeniusz Gatnar</b> , Balance of payments statistics and external competitiveness of Poland.....  | 15  |
| <b>Andrzej Sokolowski, Magdalena Czaja</b> , Efektywność metody $k$ -średnich w zależności od separowalności grup.....  | 23  |
| <b>Barbara Pawelek, Józef Pocięcha, Adam Sagan</b> , Wielosektorowa analiza ukrytych przejść w modelowaniu zagrożenia upadłością polskich przedsiębiorstw .....                                     | 30  |
| <b>Elżbieta Gołata</b> , Zróżnicowanie procesu starzenia i struktur demograficznych w Poznaniu i aglomeracji poznańskiej na tle wybranych dużych miast Polski w latach 2002-2011.....               | 39  |
| <b>Aleksandra Łuczak, Feliks Wysocki</b> , Ustalanie systemu wag dla cech w zagadnieniach porządkowania liniowego obiektów .....  | 49  |
| <b>Marek Walesiak</b> , Wzmacnianie skali pomiaru dla danych porządkowych w statystycznej analizie wielowymiarowej .....  | 60  |
| <b>Paweł Lula</b> , Identyfikacja słów i fraz kluczowych w tekstach polskojęzycznych za pomocą algorytmu <i>RAKE</i> .....  | 69  |
| <b>Mariusz Kubus</b> , Propozycja modyfikacji metody złagodzonego LASSO.....  | 77  |
| <b>Andrzej Bąk, Tomasz Bartłomowicz</b> , Wielomianowe modele logitowe wyborów dyskretnych i ich implementacja w pakiecie <i>DiscreteChoice</i> programu R.....                                     | 85  |
| <b>Justyna Brzezińska</b> , Wykorzystanie modeli logarytmiczno-liniowych do analizy bezrobocia w Polsce w latach 2004-2012.....   | 95  |
| <b>Andrzej Bąk, Marcin Pelka, Aneta Rybicka</b> , Zastosowanie pakietu <i>dcMNM</i> programu R w badaniach preferencji konsumentów wódki .....  | 104 |
| <b>Barbara Batóg, Jacek Batóg</b> , Analiza stabilności klasyfikacji polskich województw według sektorowej wydajności pracy w latach 2002-2010 .....  | 113 |
| <b>Małgorzata Markowska, Danuta Strahl</b> , Klasyfikacja europejskiej przestrzeni regionalnej ze względu na filary inteligentnego rozwoju z wykorzystaniem referencyjnego systemu granicznego..... | 121 |
| <b>Kamila Migdał-Najman, Krzysztof Najman</b> , Formalna ocena jakości odwzorowania struktury grupowej na mapie Kohonena .....  | 131 |
| <b>Kamila Migdał-Najman, Krzysztof Najman</b> , Graficzna ocena jakości odwzorowania struktury grupowej na mapie Kohonena .....   | 139 |
| <b>Beata Basiura, Anna Czapkiewicz</b> , Badanie jakości klasyfikacji szeregów czasowych .....  | 148 |
| <b>Michał Trzęsiok</b> , Wybrane metody identyfikacji obserwacji oddalonych.....  | 157 |

|   |     |
|---|-----|
| <b>Grażyna Dehnel, Tomasz Klimanek</b> , Taksonomiczne aspekty estymacji pośredniej uwzględniającej autokorelację przestrzenną w statystyce gospodarczej.....                     | 167 |
| <b>Michał Bernard Pietrzak, Justyna Wilk</b> , Odległość ekonomiczna w modelowaniu zjawisk przestrzennych z wykorzystaniem modelu grawitacji.....                                 | 177 |
| <b>Maciej Beręsewicz</b> , Próba zastosowania różnych miar odległości w uogólnionym estymatorze Petersena.....  | 186 |
| <b>Marcin Szymkowiak, Tomasz Józefowski</b> , Konstrukcja i praktyczne wykorzystanie estymatorów typu SPREE na przykładzie dwuwymiarowych tabel kontyngencji.....                 | 195 |
| <b>Marcin Pelka</b> , Klasyfikacja pojęciowa danych symbolicznych w podejściu wielomodelowym.....   | 202 |
| <b>Małgorzata Machowska-Szewczyk</b> , Ocena klas w rozmytej klasyfikacji obiektów symbolicznych.....   | 210 |
| <b>Justyna Wilk</b> , Problem wyboru liczby klas w taksonomicznej analizie danych symbolicznych.....  | 220 |
| <b>Andrzej Dudek</b> , Metody analizy skupień w klasyfikacji markerów map Google.....   | 229 |
| <b>Ewa Roszkowska</b> , Ocena ofert negocjacyjnych w słabo ustrukturyzowanych problemach negocjacyjnych z wykorzystaniem rozmytej procedury SAW.....                              | 237 |
| <b>Marcin Szymkowiak, Marek Witkowski</b> , Zastosowanie analizy korespondencji do badania kondycji finansowej banków spółdzielczych.....   | 248 |
| <b>Bartłomiej Jefmański</b> , Budowa rozmytych indeksów satysfakcji klientów z zastosowaniem programu R.....  | 257 |
| <b>Karolina Bartos</b> , Odkrywanie wzorców zachowań konsumentów za pomocą analizy koszykowej danych transakcyjnych.....  | 266 |
| <b>Joanna Trzęsiok</b> , Taksonomiczna analiza krajów pod względem dzietności kobiet oraz innych czynników demograficznych.....   | 275 |
| <b>Beata Bal-Domańska</b> , Próba identyfikacji większych skupisk regionalnych oraz ich konwergencja.....   | 285 |
| <b>Beata Bieszk-Stolorz, Iwona Markowicz</b> , Wpływ zasiłku na proces poszukiwania pracy.....  | 294 |
| <b>Marta Dziechciarz-Duda, Klaudia Przybysz</b> , Wykształcenie a potrzeby rynku pracy. Klasyfikacja absolwentów wyższych uczelni.....  | 303 |
| <b>Tomasz Klimanek</b> , Problem pomiaru procesu dezagrarnizacji wsi polskiej w świetle wielowymiarowych metod statystycznych.....  | 313 |
| <b>Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska</b> , Wybrane metody analizy danych wzdluznych.....   | 321 |
| <b>Artur Zaborski</b> , Zastosowanie miar odległości dla danych porządkowych do agregacji preferencji indywidualnych.....   | 330 |
| <b>Mariola Chrzanowska, Nina Drejerska, Iwona Pomianek</b> , Zastosowanie analizy korespondencji do badania sytuacji mieszkańców strefy podmiejskiej Warszawy na rynku pracy..... | 338 |

|   |     |
|---|-----|
| <b>Katarzyna Wawrzyniak</b> , Klasyfikacja województw według stopnia realizacji priorytetów Strategii Rozwoju Kraju 2007-2015 z wykorzystaniem wartości centrum wierszowego ..... | 346 |
|---|-----|

## Summaries

|   |     |
|---|-----|
| <b>Eugeniusz Gatnar</b> , Statystyka bilansu płatniczego a konkurencyjność gospodarki Polski .....  | 22  |
| <b>Andrzej Sokółowski, Magdalena Czaja</b> , Cluster separability and the effectiveness of $k$ -means method .....  | 29  |
| <b>Barbara Pawelek, Józef Pocięcha, Adam Sagan</b> , Multisectoral analysis of latent transitions in bankruptcy prediction models.....  | 38  |
| <b>Elżbieta Golata</b> , Differences in the process of aging and demographic structures in Poznań and the agglomeration compared to selected Polish cities in the years 2002-2011 ..... | 48  |
| <b>Aleksandra Łuczak, Feliks Wysocki</b> , Determination of weights for features in problems of linear ordering of objects .....  | 59  |
| <b>Marek Walesiak</b> , Reinforcing measurement scale for ordinal data in multivariate statistical analysis .....   | 68  |
| <b>Paweł Lula</b> , Automatic identification of keywords and keyphrases in documents written in Polish.....   | 76  |
| <b>Mariusz Kubus</b> , The proposition of modification of the relaxed LASSO method.....   | 84  |
| <b>Andrzej Bąk, Tomasz Bartłomowicz</b> , Microeconomic multinomial logit models and their implementation in the <code>DiscreteChoice</code> R package .                                | 94  |
| <b>Justyna Brzezińska</b> , The analysis of unemployment data in Poland in 2004-2012 with application of log-linear models .....  | 103 |
| <b>Andrzej Bąk, Marcin Pelka, Aneta Rybicka</b> , Application of the MMLM package of R software for vodka consumers preference analysis.....  | 112 |
| <b>Barbara Batóg, Jacek Batóg</b> , Analysis of the stability of classification of Polish voivodeships in 2002-2010 according to the sectoral labour productivity .....                 | 120 |
| <b>Małgorzata Markowska, Danuta Strahl</b> , Classification of the European regional space in terms of smart growth pillars using the reference limit system.....                       | 130 |
| <b>Kamila Migdał Najman, Krzysztof Najman</b> , Formal quality assessment of group structure mapping on the Kohonen's map .....   | 138 |
| <b>Kamila Migdał Najman, Krzysztof Najman</b> , Graphical quality assessment of group structure mapping on the Kohonen's map .....  | 147 |
| <b>Beata Basiura, Anna Czapkiewicz</b> , Validation of time series clustering .....   | 156 |
| <b>Michał Trzęsiok</b> , Selected methods for outlier detection.....  | 166 |

|  |     |
|--|-----|
| <b>Grażyna Dehnel, Tomasz Klimanek</b> , Taxonomic aspects of indirect estimation accounting for spatial correlation in enterprise statistics .....  | 176 |
| <b>Michał Bernard Pietrzak, Justyna Wilk</b> , Economic distance in modeling spatial phenomena with the application of gravity model.....  | 185 |
| <b>Maciej Beręsewicz</b> , An attempt to use different distance measures in the Generalized Petersen estimator .....   | 194 |
| <b>Marcin Szymkowiak, Tomasz Józefowski</b> , Construction and practical using of SPREE estimators for two-dimensional contingency tables.....   | 201 |
| <b>Marcin Pelka</b> , The ensemble conceptual clustering for symbolic data.....  | 209 |
| <b>Małgorzata Machowska-Szewczyk</b> , Evaluation of clusters obtained by fuzzy classification methods for symbolic objects.....   | 219 |
| <b>Justyna Wilk</b> , Problem of determining the number of clusters in taxonomic analysis of symbolic data .....   | 228 |
| <b>Andrzej Dudek</b> , Clustering techniques for Google maps markers.....  | 236 |
| <b>Ewa Roszkowska</b> , The evaluation of negotiation offers in ill structure negotiation problems with the application of fuzzy SAW procedure .....   | 247 |
| <b>Marcin Szymkowiak, Marek Witkowski</b> , The use of correspondence analysis in analysing the financial situation of cooperative banks.....  | 256 |
| <b>Bartłomiej Jefmański</b> , The construction of fuzzy customer satisfaction indexes using R program.....   | 265 |
| <b>Karolina Bartos</b> , Discovering patterns of consumer behaviour by market basket analysis of the transactional data.....   | 274 |
| <b>Joanna Trzęsiok</b> , Cluster analysis of countries with respect to fertility rate and other demographic factors .....  | 284 |
| <b>Beata Bal-Domańska</b> , An attempt to identify major regional clusters and their convergence .....   | 293 |
| <b>Beata Bieszk-Stolorz, Iwona Markowicz</b> , The influence of benefit on the job finding process .....   | 302 |
| <b>Marta Dziechciarz-Duda, Klaudia Przybysz</b> , Education and labor market needs. Classification of university graduates .....   | 312 |
| <b>Tomasz Klimanek</b> , The problem of measuring deagrarianisation process in rural areas in Poland using multivariate statistical methods.....   | 320 |
| <b>Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska</b> , Selected methods for an analysis of longitudinal data.....   | 329 |
| <b>Artur Zaborski</b> , The application of distance measures for ordinal data for aggregation individual preferences .....   | 337 |
| <b>Mariola Chrzanowska, Nina Drejerska, Iwona Pomianek</b> , Application of correspondence analysis to examine the situation of the inhabitants of Warsaw suburban area in the labour market .....                             | 345 |
| <b>Katarzyna Wawrzyniak</b> , Classification of voivodeships according to the level of the realization of priorities of <i>the National Development Strategy 2007-2015</i> with using the values of centroid of the rows ..... | 355 |

**Mariola Chrzanowska, Nina Drejerska, Iwona Pomianek**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

---

## **ZASTOSOWANIE ANALIZY KORESPONDENCJI DO BADANIA SYTUACJI MIESZKAŃCÓW STREFY PODMIEJSKIEJ WARSZAWY NA RYNKU PRACY**

---

**Streszczenie:** Strefa podmiejska Warszawy wyróżnia się na tle całego Mazowsza stosunkowo dobrą sytuacją na rynku pracy. Badania empiryczne definiujące m.in. kwestie aktywności zawodowej jej mieszkańców przeprowadzono w wybranych 30 gminach województwa mazowieckiego. Badaniem w formie wywiadu kwestionariuszowego objęto 800 gospodarstw domowych. W artykule przedstawiono częściowe wyniki badania, dotyczące sytuacji 836 mieszkańców strefy podmiejskiej na rynku pracy. Do eksploracji wyników zastosowano analizę korespondencji. Uzyskane wyniki, grupując oddzielnie lokalnych przedsiębiorców, potwierdzają znaczenie drobnej przedsiębiorczości w różnicowaniu miejsc pracy na obszarach wiejskich strefy podmiejskiej. Potrzebę dalszych badań i analiz wskazują np. wyniki grupujące oddzielnie osoby pracujące w innych wsiach strefy podmiejskiej i oddzielnie osoby dojeżdżające do pracy do Warszawy.

**Słowa kluczowe:** analiza korespondencji, obszary podmiejskie, rynek pracy, Warszawa.

### **1. Wstęp**

Kryteria związane z rynkiem pracy (np. struktura zatrudnienia, dojazd do pracy do miasta centralnego) stosowane są w Polsce od lat 60. XX wieku jako podstawa wyznaczenia zasięgu aglomeracji miejskiej [Smętkowski, Jałowiecki, Gorzelak 2009]. Liczba przyjeżdżających do pracy najmniej do rdzenia na 1000 mieszkańców w wieku produkcyjnym jest jedną z siedmiu cech delimitujących Miejskie Obszary Funkcjonalne w badaniu P. Śleszyńskiego [2013]. Inni autorzy [Rosik, Stępnik, Wiśniewski 2010] porównywali trzy różne metody zbierania danych dotyczących dojazdów do pracy: Warszawskie Badania Ruchu (metoda I), dwuetapowe badanie ankietowe w ośrodku skupiającym dojazdy (Białystok) i w miejscu źródłowym tych dojazdów (19 gmin w województwie podlaskim) jako metoda II oraz badanie przepływów ludności związanych z zatrudnieniem Urzędu Statystycznego w Poznaniu na podstawie danych z bazy POLTAX (metoda III). Staty-

styki rynku pracy są również kryteriami delimitacji wykorzystywanymi na potrzeby planistyczne, np. przez Mazowieckie Biuro Planowania Regionalnego [2011].

Znaczenie dojazdów do pracy w funkcjonalnym obszarze Warszawy potwierdza przegląd badań, jak również analiza danych statystycznych. Wyniki badania dojazdów do pracy na podstawie POLTAX wskazują, że m. st. Warszawa jest na pierwszym miejscu w województwie mazowieckim pod względem liczby przyjeżdżających do pracy. Iloraz przepływu zatrudnionych wynosił tu 13,07<sup>1</sup>, czyli na jednego wyjeżdżającego do pracy z Warszawy przypadało przeciętnie około 13 przyjeżdżających do stolicy pracowników najemnych [Kruszka (red.) 2010].

Warszawski rynek pracy wskazywany jest jako największy lokalny rynek w Polsce. Wyniki analiz [Gruchociak 2012] pokazują, że składa się on z blisko 140 gmin, a jego wielkość wynika z silnego oddziaływania Warszawy na otaczające ją gminy. Według szacunku przeprowadzonego na potrzeby projektu pt. „Trendy rozwojowe Mazowsza” dominacja Warszawy jest znacznie większa, niż wynika to z rejestracji GUS. Stolica koncentruje blisko 1,3 mln miejsc pracy, a jej strefa podmiejska i przedmiejska kolejnych 400 tys. [Śleszyński (red.) 2012].

Wyniki badań własnych w strefie podmiejskiej Warszawy potwierdzają tę dominującą pozycję stolicy jako miejsca pracy mieszkańców ościennych gmin. Na około 1100 pracowników najemnych, o których uzyskano informacje z przeprowadzonych badań ankietowych, 36% pracowało w Warszawie, a tylko 19% w miejscowości zamieszkania.

Badania literaturowe oraz śledzenie bieżących dyskusji na forum instytucji publicznych wskazują na techniczne traktowanie aspektów rynku pracy. W podejściach tych najczęściej dostępne wskaźniki służą do przestrzennego wydzielenia różnych kategorii obszarów. W tej sytuacji celowe wydają się badania empiryczne głębiej wnikające w naturę procesów zachodzących na rynku pracy na wyodrębnionych obszarach funkcjonalnych. Celem niniejszego opracowania było zbadanie na podstawie dostępnego materiału empirycznego, czy wybrane charakterystyki miejsc pracy respondentów pracujących w stolicy wyróżniają się na tle pracujących poza Warszawą.

## 2. Materiał badawczy

Badanie przeprowadzono w ramach projektu NCN pt. „Ekonomiczno-społeczne uwarunkowania rozwoju obszarów wiejskich województwa mazowieckiego stanowiących strefę podmiejską i zewnętrzną Warszawy”, N N114 145240. Punktem

---

<sup>1</sup> Na liście powiatów z największą przewagą liczby przyjeżdżających do pracy nad liczbą wyjeżdżających do pracy Warszawa zajmowała w 2006 r. pierwsze miejsce. W kolejnym powiecie krapkowskim (przemysł wydobywczy cementowo-wapienniczy) iloraz przepływów wyniósł 10,51, podczas gdy w kolejnych ośmiu miastach (powiatach grodzkich) w pierwszej dziesiątce rankingu zmniejszał się od 8,91 do 5,52.



wyjścia do szczegółowego określenia obszaru badawczego były wyniki delimitacji strefy podmiejskiej i zewnętrznej Warszawy, opracowane przez P. Korcellego i P. Śleszyńskiego [Korcelli i Śleszyński 2011] oraz J. Bańskiego [Bański (red.) 2009]. Po zestawieniu i przeanalizowaniu tych dwóch klasyfikacji wybrano do badań: gminy wiejskie według definicji GUS (wg kryterium administracyjnego) oraz gminy miejsko-wiejskie z miastami o liczbie ludności nieprzekraczającej 5 tys. W konsekwencji otrzymano 26 gmin wiejskich i 4 miejsko-wiejskie (rys. 1).



**Rys. 1.** Gminy wiejskie i miejsko-wiejskie objęte badaniem

Źródło: opracowanie własne.

Badaniem empirycznym na przełomie lat 2011 i 2012 objęto członków 800 gospodarstw domowych zamieszkujących w wybranych gminach<sup>2</sup>. Badanie przeprowadzono w formie wywiadu kwestionariuszowego. W trakcie wywiadu respondenci odpowiadali łącznie na 40 pytań. Dotyczyły one zarówno informacji o gospodarstwie domowym jak i sytuacji na rynku pracy. Dla celów niniejszego opracowania skupiono się na wybranych informacjach związanych z zatrudnieniem.

### 3. Opis badania

W pierwszym etapie z pełnej bazy danych wyodrębniono 836 osób pracujących. Spośród wszystkich respondentów w dalszych badaniach uwzględniono przedsiębiorców i pracowników najemnych. Ze względu na cel artykułu jako kryterium badawcze przyjęto pracę w miejscu zamieszkania. Jest to zmienna dychotomiczna, której warianty (tak/nie) opisano na skali nominalnej. W badaniu uwzględniono również następujące informacje:

---

<sup>2</sup> Ze względu na ograniczoną liczbę stron w niniejszym opracowaniu przedstawiono jedynie fragment badań. Pełny opis przeprowadzonych analiz wraz z procedurą doboru jednostek do badania przedstawiono w pracy [Drejerska, Chrzanowska, Pomianek 2014].

- miejsce pracy (Warszawa, inne miasto, inna wieś, zagranica, nigdzie nie dojeżdża, zmienna lokalizacja – konfiguracje poprzednich opcji, np. usługi budowlane świadczone w różnych miejscach),
- forma zatrudnienia (umowa o pracę na czas nieokreślony; umowa o pracę na czas określony; umowa o dzieło; umowa zlecenie; inne),
- wymiar czasu pracy (pełny etat; część etatu),
- sposób znalezienia obecnej pracy (ogłoszenie w prasie; bezpośrednio u pracodawcy; przez Urząd Pracy; przez znajomych/rodzinę; inne – staże/praktyki, internet),
- odległość od domu do miejsca pracy (do 10 km; 11-20 km; 21-30 km; 31 km i więcej; nie dotyczy),
- wiek ankietowanego (do 29 lat; 30-39 lat; 40-49 lat; 50-59 lat; 60-69 lat, 61 lat i więcej),
- czas dojazdu (do 20 minut, 21-40 minut, 41-60 minut, 61 minut i więcej),
- dojazd do pracy (codziennie; niecodziennie – ale co najmniej raz w tygodniu; rzadziej niż raz w tygodniu; zmienna częstotliwość; nie dojeżdża – praca w miejscu zamieszkania).

Podczas wstępnej analizy danych, okazało się, że zmienne: forma zatrudnienia oraz wymiar czasu pracy są quasi-stałe. W pierwszym przypadku dominowała umowa o pracę na czas nieokreślony (95%), natomiast w drugim około 98% respondentów pracowało w pełnym wymiarze godzin. Dlatego nie uwzględniono tych informacji w dalszych analizach.

Analiza korespondencji jest jedną z metod wielowymiarowej analizy porównawczej. Umożliwia ona trafne rozpoznanie współwystępowania kategorii zmiennych opisanych na skali nominalnej. Metoda ta jest szczegółowo opisana zarówno w literaturze polskojęzycznej [Stanimir 2005], jak i anglojęzycznej [Greenacre 2007].

Punktem wyjścia do analizy korespondencji jest sprawdzenie, czy pytania z ankiety są współzależne z punktu widzenia badanego kryterium. Z uwagi na fakt, iż większość pytań jest oparta na skali nominalnej i porządkowej do pomiaru współzależności wykorzystano test niezależności chi kwadrat. W pierwszym etapie badań sprawdzono zależność pomiędzy pracą w miejscu zamieszkania a pozostałymi zmiennymi. Wartości statystyki empirycznej  $\chi^2$  wraz z poziomem prawdopodobieństwa  $p$  odrzucenia hipotezy zerowej przedstawiono w tabeli 1.

#### 4. Wyniki badań

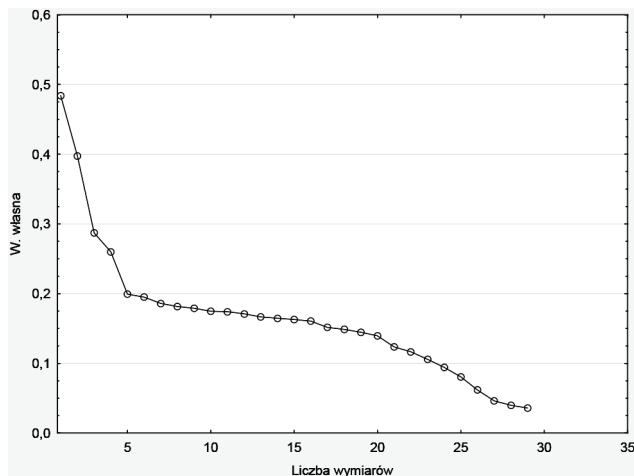
Analizę korespondencji przeprowadzono na podstawie macierzy Burta o wymiarach  $33 \times 33$ . Macierz ta została zbudowana na podstawie informacji z pozostałych sześciu zmiennych opisanych w poprzednim etapie badania. Podczas analizy wstępnej sprawdzono, jaki wymiar rzeczywistej przestrzeni współwystępowania odpowiedzi na analizowane pytania (spośród analizowanych 29 wymiarów) po-

zwolni na dokładną analizę graficzną. Posługując się kryterium łokcia (por. rys. 2), zdecydowano się na trójwymiarowy rysunek (por. rys. 3). Analizując wykres wartości własnych, należało wybrać przestrzeń pięciowymiarową. Niestety, nie jest możliwe czytelne przedstawienie takiej przestrzeni na wykresie. Dlatego zdecydowano się na przestrzeń trójwymiarową.

**Tabela 1.** Wartości statystyki  $\chi^2$  pomiędzy zmienną praca w miejscu zamieszkania a pozostałymi zmiennymi uwzględnionymi w dalszych analizach

| Zmienna                                   | Statystyka $\chi^2$ | Liczba stopni swobody | Poziom prawdopodobieństwa $p$ |
|---|---------------------|-----------------------|-------------------------------|
| Miejsce wykonywania pracy                 | 104,32              | 5                     | 0,00                          |
| Sposób znalezienia obecnej pracy          | 112,32              | 4                     | 0,00                          |
| Odległość od domu do miejsca pracy (w km) | 61,52               | 4                     | 0,00                          |
| Czas dojazdu (w minutach)                 | 46,15               | 3                     | 0,00                          |
| Wiek ankietowanego                        | 105,41              | 5                     | 0,00                          |
| Dojazd do pracy                           | 56,85               | 3                     | 0,00                          |

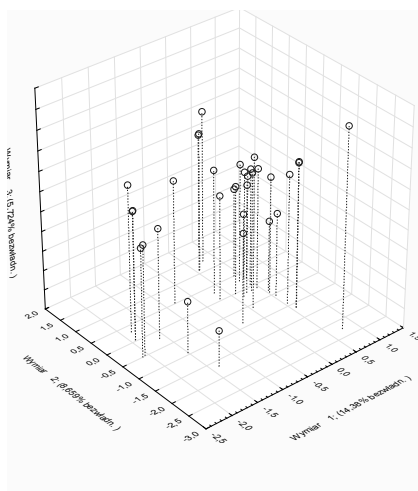
Źródło: obliczenia własne.



**Rys. 2.** Wykres wartości własnych – kryterium łokcia

Źródło: opracowanie własne.

Niestety, interpretacja wyników zaprezentowanych w przestrzeni trójwymiarowej (rys. 3) jest bardzo trudna ze względu na dużą liczbę analizowanych zmiennych i ich wariantów. Do interpretacji wyników wykorzystano zatem metodę Warda, która umożliwia wyznaczenie powiązań pomiędzy wariantami zmiennych. Do klasyfikacji metodą Warda jako zmienne wykorzystano współrzędne punktów przedstawionych na rys. 3.



**Rys. 3.** Trójwymiarowa prezentacja wyników analizy korespondencji dla analizowanych zmiennych  
 Źródło: opracowanie własne.

Analiza dendrogramu zbudowanego metodą Warda pozwoliła wyodrębnić pięć grup (rys. 4):

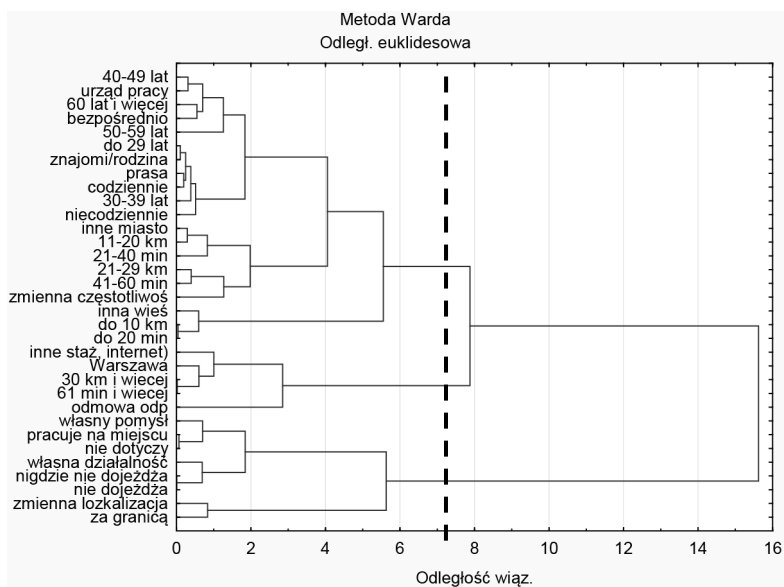
1. nieliczni członkowie badanych gospodarstw domowych pracujący za granicą oraz osoby o zmiennej lokalizacji miejsca pracy, np. wspomniane wcześniej usługi budowlane; taki wynik sugeruje potrzebę głębszej analizy wyników ankiet, uwzględniającą charakter wykonywanej pracy – czy w badanej populacji czynnikiem łączącym te dwie grupy jest rodzaj wykonywanej pracy;

2. osoby prowadzące własną działalność gospodarczą w miejscu zamieszkania; dokładniejsza analiza tej grupy wskazuje, że są to mikro- i małe firmy o lokalnej skali działalności, jak większość mikro- i małych przedsiębiorstw w Polsce;

3. pracujący w Warszawie, dojeżdżający do pracy 30 km i więcej, w czasie 60 min i dłużej, którzy znaleźli pracę np. przez Internet bądź wcześniej odbywając staż u pracodawcy; w tej grupie znajdują się również osoby, które odmówiły odpowiedzi na temat sposobu znalezienia pracy;

4. osoby pracujące w innej wsi oddalonej nie więcej niż 10 km od miejsca zamieszkania; w tym przypadku polem do dalszych badań mogłaby być próba identyfikacji czy istnieje zależność pomiędzy wiekiem, poziomem wykształcenia a przestrzenną odległością do pracy i mobilnością pod względem miejsca zatrudnienia;

5. ostatnia grupa charakteryzuje się największym zróżnicowaniem, są w niej osoby zarówno z górnych, jak i dolnych przedziałów wiekowych, o różnych sposobach znalezienia zatrudnienia i różnych odległościach pomiędzy miejscem zamieszkania a wykonywania pracy; dość jednolitą podgrupą wydają się tu osoby w wieku 40+, które znalazły zatrudnienie w tradycyjny sposób bądź w bezpośrednim kontakcie z pracodawcą czy też za pośrednictwem urzędu pracy.



Rys. 4. Diagram hierarchicznej klasyfikacji przeprowadzonej metodą Warda

Źródło: opracowanie własne.

## 5. Podsumowanie

Quasi-stałość zmiennych: forma zatrudnienia i wymiar czasu pracy, wskazuje na słabo rozwinięty rynek pracy. Należy to podkreślić w kontekście badania mieszkańców gmin strefy podmiejskiej Warszawy, a więc obszaru postrzeganego tradycyjnie jako miejsce o dobrej sytuacji na rynku pracy, co więcej uważanego za względnie rozwinięty i zróżnicowany rynek pracy.

Wyniki przeprowadzonej analizy korespondencji wskazują na takie grupowania mieszkańców obszarów podmiejskich Warszawy<sup>3</sup>, które pozwalają na zidentyfikowanie ważnych tendencji na obszarze funkcjonalnym największego miasta w Polsce. Jedną z ważniejszych konkluzji wydaje się wyodrębnienie grupy przedsiębiorców, właścicieli mikro- i małych firm, prowadzących działalność na obszarach wiejskich strefy podmiejskiej. Potwierdza to rolę małej przedsiębiorczości w różnicowaniu miejsc zatrudnienia także w okolicach dużych miast, podczas gdy tradycyjnie znaczenie drobnej przedsiębiorczości jest wskazywane jako bardzo ważne dla obszarów typowo wiejskich. Wyniki wskazujące na pracę w innych wsiach strefy podmiejskiej czy dojazd do pracy do Warszawy stawiają pytania o determinanty takiej sytuacji. Zagadnienia te wyznaczają kierunki dalszych badań.

<sup>3</sup> Są to: pracujący za granicą i w różnych lokalizacjach; prowadzący własną działalność gospodarczą z miejsca zamieszkania, dojeżdżający do pracy w Warszawie, osoby w wieku produkcyjnym niemobilnym pracujące blisko miejsca zamieszkania.

## Literatura

- Bański J. (red.) (2009), *Analiza zróżnicowania i perspektywy rozwoju obszarów wiejskich w Polsce do 2015 roku*, Studia Obszarów Wiejskich, t. 16, PTG, IGIPZ PAN, Warszawa.
- Drejerska N., Chrzanowska M., Pomianek I., *Strefa podmiejska Warszawy. Wybrane zagadnienia*, <http://kpefpim.wne.sggw.pl/badania/projekt-ncn-n-n114-145240/> (15 lutego 2014).
- Greenacre M. (2007), *Correspondence Analysis in Practice*, Chapman & Hall/CR, London.
- Gruchocki H. (2012), *Delimitacja lokalnych rynków pracy w Polsce. Przegląd statystyczny*, numer specjalny 2, Warszawa, s. 277-297.
- Korcelli P., Śleszyński P. (2011), *Mapa 4. Funkcjonalne obszary miejskie. Koncepcja przestrzennego zagospodarowania kraju 2030. Projekt dokumentu rządowego przeznaczony do konsultacji*, Ministerstwo Rozwoju Regionalnego, Warszawa.
- Kruszka K. (red.) (2010), *Dojazdy do pracy w Polsce. Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem*, Główny Urząd Statystyczny, Urząd Statystyczny w Poznaniu, Poznań, s. 46.
- Mazowieckie Biuro Planowania Regionalnego (2011), *Studium planu zagospodarowania przestrzennego obszaru metropolitalnego Warszawy*, Warszawa.
- Rosik P., Stępnik M., Wiśniewski R. (2010), *Dojazdy do pracy do Warszawy i Białegostoku*, Studia Regionalne i Lokalne, nr 2(40), s. 77-98.
- Smętkowski M., Jałowecki B., Gorzelak G. (2009), *Obszary metropolitalne w Polsce: problemy rozwojowe i delimitacja*, Raporty i analizy EUROREG 1, Aneks 3, Warszawa.
- Stanimir A. (2005), *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wyd. Akademii Ekonomicznej, Wrocław.
- Śleszyński P. (2013), *Delimitacja miejskich obszarów funkcjonalnych stolic województw*, „Przegląd Geograficzny” 85, 2, s. 173-197.
- Śleszyński P. (red.) (2012), *Społeczno-demograficzne uwarunkowania rozwoju Mazowsza*, „Trendy Rozwojowe Mazowsza” nr 3, Warszawa, s. 29.

### APPLICATION OF CORRESPONDENCE ANALYSIS TO EXAMINE THE SITUATION OF THE INHABITANTS OF WARSAW SUBURBAN AREA IN THE LABOUR MARKET

**Summary:** The suburban zone of Warsaw with its relatively good situation in the labor market stands out from the whole Mazovia. Empirical research concerning economic activity of this zone inhabitants was carried out in 30 selected municipalities of Mazowieckie voivodeship. The study included 800 households and was conducted in the form of a questionnaire interview. This paper presents partial results on the situation of the 836 inhabitants of the suburban area in the labor market. To explore the results the correspondence analysis was used. The results obtained by grouping local entrepreneurs separately confirm the importance of small business in the differentiation of jobs in rural areas of the suburban zone. The results grouping separately those working in other villages of the suburban area and those commuting to Warsaw indicate the need for further research and analysis.

**Keywords:** correspondence analysis, suburban area, labour market, Warsaw.