

# 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

**Karolina Paradysz**

Centrum Statystyki Regionalnej – Uniwersytet Ekonomiczny w Poznaniu

---

## **HIERARCHICZNA METODA GRUPOWANIA POWIATÓW JAKO PODEJŚCIE BENCHMARKOWE W OCENIE BEZROBOCIA WEDŁUG BAEL-U W WYBRANYCH TYPACH MAŁYCH OBSZARÓW**

---

**Streszczenie:** Jednym z rodzajów oceny dobroci wyników estymacji dla małych obszarów jest kryterium merytoryczne [Paradysz 2008]. W przypadku estymacji wskaźników bezrobocia dla powiatów merytorycznym kryterium oceny (benchmarkiem) mogą być pewne rozpoznane wcześniej prawidłowości na lokalnych rynkach pracy: aglomeracje wielkomiejskie, powiaty przygraniczne, tereny turystyczne. Wykorzystując hierarchiczną metodę grupowania, dokonano merytorycznej oceny jakości estymacji stopy bezrobocia w przekroju powiatowym, które wykonano wcześniej w GUS [Bracha i in. 2004]. Wyniki tylko częściowo potwierdzają przyjęte założenia, co oznacza jakość estymacji dla małych obszarów nie była wówczas wystarczająco dobra.

**Słowa kluczowe:** statystyka małych obszarów, benchmarking, BAEL, rynek pracy.

### **1. Wstęp**

Problemem badawczym jest ocena jakości estymacji w statystyce małych obszarów (SMO) [por. Rao 2003] w badaniu rynku pracy dokonanym przez zespół w Głównym Urzędzie Statystycznym. Jest to kontynuacja artykułu [Paradysz 2013], w którym oceniono jakość estymatorów z zakresu SMO z punktu widzenia kryterium formalnego poziomu i kolejności [Paradysz 2008]. Prezentowana tutaj ocena jakości estymatorów nawiązuje do kryterium merytorycznego także zaproponowanego przez J. Paradysza. W niniejszym artykule jako jedno z tych kryteriów zastosowano merytoryczną ocenę wyników estymacji pośredniej. Merytoryczne kryterium oceny tych wyników polega na przykład na tym, by zostały zachowane: logiczny układ współczynników i zgodność z dotychczasowym stanem wiedzy. Dla obszarów – bez względu na liczbę wylosowanych jednostek, tzn. dla wszystkich powiatów – wskaźniki precyzji i obciążenia zostaną poddane ocenie w dwóch ujęciach: zbadane będą zachowania tych wskaźników w powiatach przygranicznych



(celem porównania ze wskaźnikami w powiatach w centrum Polski) oraz w powiatach okalających duże miasta w Polsce. W artykule zaproponowano nowe zastosowanie dobrze znanej metody, a mianowicie hierarchicznej metody grupowania. Metoda ta z powodzeniem stosowana jest w przypadkach, gdy chcemy odpowiedzieć na pytanie, jakie jest podobieństwo badanych obiektów na podstawie określonych wskaźników [Ryś-Jurek 2012]. W tradycyjnym ujęciu w analizie taksonomicznej na podstawie przyjętych *a priori* zmiennych diagnostycznych szuka się podobieństwa między obiektami, w tym przypadku są to powiaty grodzkie i ziemskie. Wyselekcjonowane do analizy zmienne traktuje się najczęściej jako pewne źródło informacji, które dobrze opisuje dane objekty.

Propozycja zawarta w niniejszej pracy polega na odwróceniu sytuacji. Przyjęte zostało bowiem założenie, że dobrze znamy uporządkowanie obiektów na podstawie wcześniejszej wiedzy, na przykład uporządkowanie powiatów według bezrobocia rejestrowanego lub innych wskaźników skorelowanych z ogólną sytuacją ekonomiczną danego regionu, i potraktować to jako benchmark dla estymatorów klasy SMO. Innymi słowy, przyjmując wyniki estymacji C. Brachy, B. Lednickiego i R. Wieczorkowskiego [2004] jako zmienne diagnostyczne w analizie taksonomicznej i stosując metodę hierarchiczną grupowania, dokonano podziału powiatów grodzkich.

## 2. Podział regionów pod względem bezrobocia

Przed przystąpieniem do analizy taksonomicznej scharakteryzuję regiony wyróżniające się pod względem poziomu bezrobocia. Z tego punktu widzenia, na podstawie literatury przedmiotu [Smeńkowski i in. 2009; GUS 2013], wśród powiatów grodzkich i ziemskich można wyróżnić między innymi obszary metropolitalne, turystyczne i przygraniczne.

### 2.1. Specyfika lokalnych rynków pracy na obszarach metropolitalnych

Z punktu widzenia badań nad rynkiem pracy obszary metropolitalne mają swoją specyfikę, należą do niej:

- koncentracja przedsiębiorczości i miejsc pracy, w tym miejsc o najwyższej jakości życia, na obszarach znajdujących się w sieci międzynarodowych powiązań;
- wzrost różnic między centrami i peryferiami i oraz częstotliwości i gęstości poziomych relacji między metropoliami;
- powiększanie się nierówności między regionami i miastami oraz wewnątrz stref metropolitalnych.

Ośrodki miejskie w Polsce nie są w pełni równomiernie rozmieszczone. Zachodnie obszary kraju charakteryzują się większą gęstością miast niż wschodnie,

co wynika z historycznych procesów dyfuzji urbanizacji [Gołata 2004]. Współcześnie najwięcej miast jest w województwach wielkopolskim i dolnośląskim, najmniej zaś na obszarze województw podlaskiego i lubelskiego [Smętkowski i in. 2009, s. 55].

## **2.2. Specyfika lokalnych rynków pracy w miejscowościach turystycznych**

Rynek pracy charakteryzuje się pewną sezonowością w zależności od pory roku oraz terytorium. Kształtowanie się wielkości charakterystycznych dla rynku pracy w kontekście sezonowości wiąże się ściśle ze wspomnianym wcześniej kryterium merytorycznym określonym przez Paradysza [Paradysz 2008]. Sezonowość pracy i turystyka zdają się pojęciami nierozzerwalnymi. Również migracje związane z poszukiwaniem pracy sezonowej wiążą się z określonym obszarem kraju. Na przykład miejscowości nadmorskie charakteryzują się sezonowością pracy w miesiącach letnich, natomiast miejscowości górskie generują miejsca pracy w miesiącach zimowych. Sezonowość migracji związanych z pracą koreluje z migracjami ludności w celach turystycznych.

## **2.3. Specyfika lokalnych rynków pracy na obszarach przygranicznych**

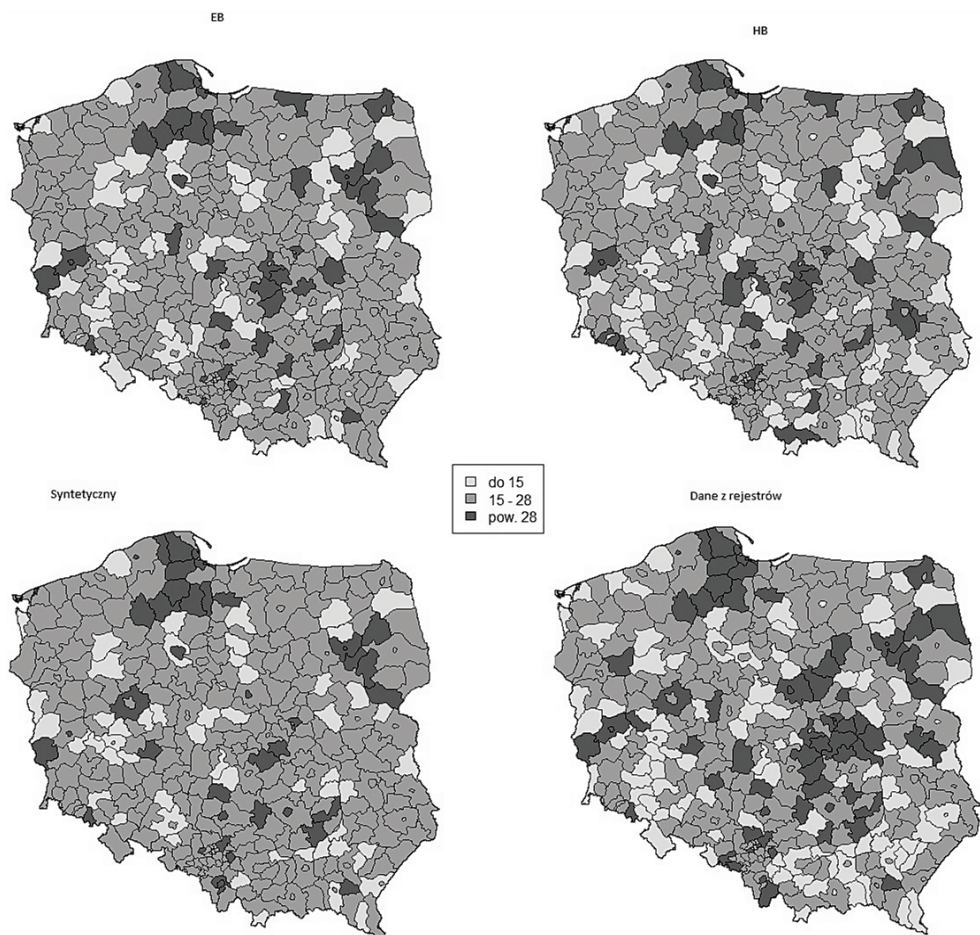
W momencie wstępowania Polski do Unii Europejskiej wpływ migracji na wskaźniki rynku pracy – wbrew powszechnym przekonaniom – był umiarkowany. Wyjazdy Polaków za granicę miały natomiast daleko idące konsekwencje dla struktury polskiego rynku pracy i spójności społecznej. Owe konsekwencje objawiły się w tak różnorodnych zjawiskach, jak presja płacowa, nowe wzorce karier zawodowych czy poprawa sytuacji dochodowej polskich rodzin. Jak to określa Urząd Komitetu Integracji Europejskiej: „Lata 2004-2007 to okres bezprecedensowego wzrostu skali mobilności Polaków – liczba osób podejmujących pracę w krajach UE wzrosła w tym czasie z około 1 mln osób do 2,3 mln” [Kałużyńska i in. (red.) 2009, s. 16].

## **3. Zastosowanie hierarchicznej metody grupowania oraz prezentacja wyników**

### **3.1. Hierarchiczna metoda grupowania**

W analitycznej części artykułu, uwzględniającego specyfikę rynku pracy, wzięto pod uwagę stopę bezrobocia na szczeblu lokalnym, czyli NTS 4 – powiaty. W porównaniach uwzględniono jedynie trzy estymatory: syntetyczny, hierarchiczno-bayesowski (HB) oraz empiryczno-bayesowski (EB). Są to estymatory wykorzystane przez zespół pod kierunkiem C. Brachy, oszacowane dla powiatów (dla każdego z osobna) w całym kraju [Bracha 2004]. Pozostałe zaś szacunki (estymatory

klasyczny i złożony) zostały wykonane tylko dla niektórych jednostek terytorialnych<sup>1</sup>. Jako benchmark (wzorzec – podstawa porównań) wzięto pod uwagę dane o bezrobociu z rejestrów administracyjnych (bezrobocie rejestrowane). Stopę bezrobocia (opublikowaną przez GUS) dla 2003 r., w celu zachowania porównywalności danych, zaprezentowano w dwóch wariantach różniących się źródłem danych o pracujących w rolnictwie indywidualnym, będących składową cywilnej ludności aktywnej zawodowo. Dane prezentowane w wariantcie A opracowane zostały



**Rys. 1.** Stopy bezrobocia rejestrowanego w układzie powiatów oraz według estymatorów klasy SMO (EB, HB, syntetyczny) w 2003 r.

Źródło: opracowanie własne na podstawie [Bracha i in. 2004; GUS 2013].

<sup>1</sup> Wartości estymatorów klasycznego i złożonego zostały wyznaczone dla powiatów i miast na prawach powiatu, w których wylosowano powyżej 50 jednostek.

z wykorzystaniem wyników Powszechnego Spisu Rolnego, przeprowadzonego w 1996 r., natomiast w wariancie B wykorzystano wyniki Narodowego Spisu Powszechnego Ludności i Mieszkań przeprowadzonego w 2002 r. oraz Powszechnego Spisu Rolnego z 2002 r. W analizie wykorzystano wariant A. Dane zostały przedstawione na kartogramach (rys. 1).

W każdym z czterech przypadków wysokość stopy bezrobocia na terenach przygranicznych nie odbiega znacząco od wielkości tejże zmiennej w centrum terytorium kraju. Ciemniejsze kolory, świadczące o podwyższonej stopie bezrobocia, dostrzec można w przypadku północnej Polski (tereny Trójmiasta), co dowodzi sezonowego charakteru pracy w miejscowościach nadmorskich.

Do badania dużych aglomeracji wybrano 17 miast, dla których zespół Brachy wykonał 5 estymatorów – do trzech doszły jeszcze klasyczny i złożony. Jako benchmark wykorzystano stopę bezrobocia rejestrowanego z Banku Danych Lokalnych [por. GUS 2013]. Najpierw wyznaczono wzorzec (benchmark), czyli punkt odniesienia np. danych pochodzących z innych źródeł, a następnie – odległości każdego obiektu od wzorca (benchmarku). Głównym celem takiego postępowania jest odpowiedź na pytanie, jak bardzo pewne obiekty są do siebie podobne, a w szczególności: w jakim stopniu podobne są do wzorca [Paradysz, Paradysz 2011]. Im bardziej obiekty są podobne, tym są sobie bliższe, czyli odległość między nimi jest mniejsza. Wyróżnić można benchmarking wewnętrzny (porównanie estymatorów między sobą) i zewnętrzny (porównanie z wzorcem – dane z rejestru administracyjnego – bezrobocie rejestrowane). W tym celu wykorzystano metodę Warda.

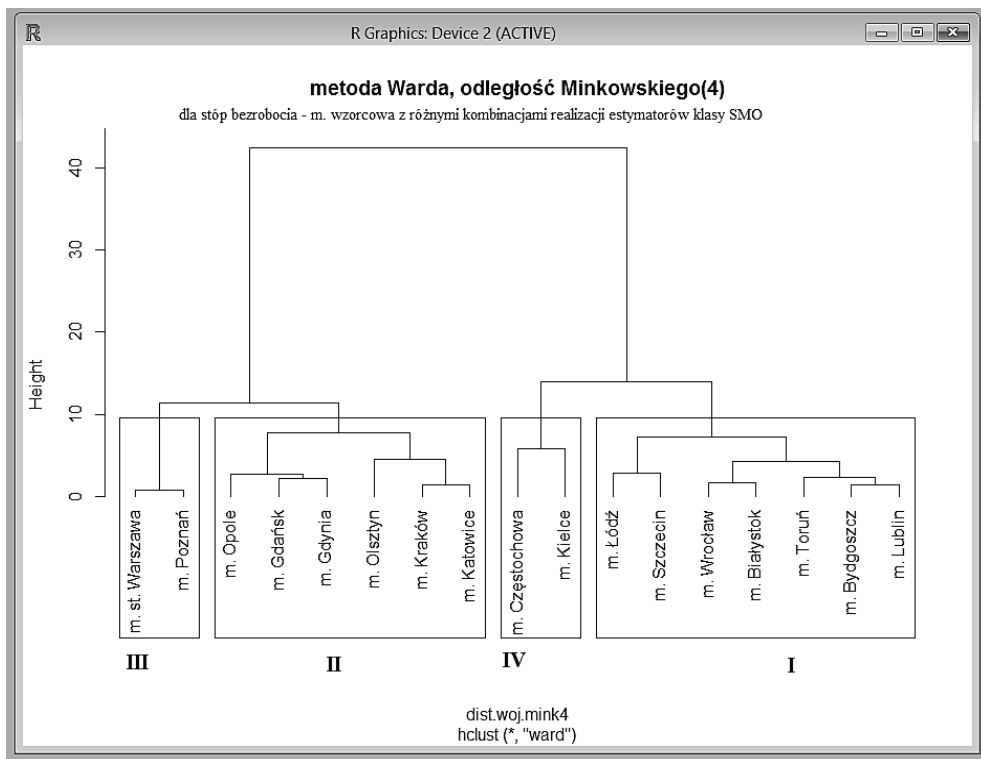
W powszechnie stosowanym podejściu do analizy rozwoju regionalnego na podstawie przyjętych wskaźników rozwoju dokonuje się oceny obiektów, w tym obszarów, poprzez szukanie ich podobieństwa. W tej pracy cel badawczy został odwrócony i na podstawie otrzymanej klasyfikacji małych obszarów oceniono wskaźniki, którymi są poszczególne estymatory stóp bezrobocia. Innymi słowy, mniej lub bardziej logiczna struktura obszarów metropolitalnych posłuży do pośredniej oceny danego estymatora. Możliwe są następujące cztery podejścia: metoda bezwzorcowa – każdy estymator osobno, metoda bezwzorcowa – różne kombinacje realizacji estymatorów klasy SMO<sup>2</sup>, metoda wzorcowa – każdy estymator osobno, metoda wzorcowa – różne kombinacje realizacji estymatorów klasy SMO. Ze względu na ograniczoną objętość artykułu przedstawione zostanie ostatnie podejście.

Pierwsza grupa to miasta, w których stopa bezrobocia zarówno w przypadku benchmarku (danych z rejestru), jak i w przypadku estymatorów jest dość wysoka. W drugiej grupie znajdują się miasta, w których różnice między stopami bezrobocia w przypadku benchmarku i estymatorów nie są stosunkowo duże (w porównaniu z pozostałymi grupami). Trzecia grupa charakteryzuje się najniższymi stopami

---

<sup>2</sup> W klasycznym podejściu do SMO odpowiadałoby to estymatorom złożonym.

bezrobocia zarówno w przypadku wzorca, jak i estymatorów. Uzyskany obraz bezrobocia w dużych miastach na podstawie analizy taksonomicznej, dla której punktem wyjścia były dane ze statystyki małych obszarów, potwierdza wcześniejsze konstatacje innych autorów. Zatem wydaje się, że w przypadku większych miast estymacja wskaźników bezrobocia za pomocą estymatorów pośrednich była zasadniczo dobra.



**Rys. 2.** Dendrogram podobieństwa dla stóp bezrobocia – metoda wzorcowa z różnymi kombinacjami realizacji estymatorów klasy SMO

Źródło: opracowanie własne na podstawie [Bracha i in. 2004; GUS 2013].

W tabeli 1 pokazano kształtowanie się stóp bezrobocia uzyskanych w toku estymacji. Najbardziej istotnym wnioskiem, wynikającym zarówno z diagramu, jak i tabeli, jest fakt, iż uwzględnione w badaniu miasta podzielone zostały (przy zastosowaniu hierarchicznej metody grupowania) na grupy zgodnie ze znanym logicznym uporządkowaniem. Można zatem stwierdzić, że takie właśnie uporządkowanie skorelowane jest z ogólną sytuacją ekonomiczną danego regionu. Na przykład w grupie trzeciej znalazły się miasta o stosunkowo niskich stopach bezrobocia, a w czwartej – o stosunkowo wysokiej stopie.

**Tabela 1.** Wyniki zastosowania hierarchicznej metody grupowania

Miasto	Stopa bezrobocia:					
	rejestrowanego	estymator HB	estymator EB	estymator klasyczny	estymator syntetyczny	estymator złożony
grupa I						
Wrocław	12,8	20,1	20,5	20,6	18,3	19,5
Bydgoszcz	12,8	18,2	18,5	18,4	17,2	17,8
Toruń	15,1	18,2	19,1	19,3	18,6	19,0
Lublin	13,4	19,2	18,5	19,0	18,4	18,7
Łódź	19,0	20,0	20,3	20,4	20,0	20,2
Białystok	13,9	20,3	21,0	21,0	19,9	20,4
Szczecin	16,2	20,2	20,6	20,9	19,0	19,9
grupa II						
Kraków	8,3	14,7	15,1	15,3	17,9	16,7
Opole	10,9	11,8	12,6	12,1	13,5	12,8
Gdańsk	12,7	13,1	13,0	12,9	15,9	14,5
Gdynia	10,5	13,2	13,2	13,7	15,5	14,5
Katowice	8,4	15,6	16,4	16,2	17,1	16,7
Olsztyn	11,8	16,1	15,5	15,2	15,0	15,1
grupa III						
Warszawa	6,3	11,8	11,8	11,8	12,9	12,4
Poznań	7,1	12,2	12,0	11,8	12,5	12,1
grupa IV						
Częstochowa	16,2	12,4	19,2	18,9	22,7	20,7
Kielce	15,5	13,8	14,2	14,5	20,0	17,5

Źródło: opracowanie własne na podstawie [Bracha i in. 2004; GUS 2013].

#### 4. Wnioski

Na podstawie merytorycznej oceny wyników estymacji pośredniej stwierdzono, że układ przestrzenny wskaźników bezrobocia jest logiczny i zgodny z dotychczasowym stanem wiedzy:

- mapy stóp bezrobocia w powiatach całego kraju w każdym z czterech przypadków (estymatorów: HB, EB, syntetycznego) oraz benchmarku, czyli bezrobocie rejestrowane, w tym na terenach przygranicznych, są zasadniczo podobne;
- stopa bezrobocia w całej Polsce charakteryzuje się sezonowością na obszarach turystycznych (zarówno w przypadku wielkości szacunkowych HB, EB i syntetycznych, jak i tych z rejestrów administracyjnych – zgodność benchmarkowa);
- podział na grupy uzyskany w toku zaproponowanej zmienionej metody podobieństwa w pewnym przybliżeniu potwierdził nasze przypuszczenia co do przynależności do danej grupy.

## Literatura

- Bracha C., Lednicki B., Wieczorkowski R. (2004), *Wykorzystanie złożonych metod estymacji do dezagregacji danych z badania aktywności ekonomicznej ludności w roku 2003*, Z Prac Zakładu Badań Statystyczno-Ekonomicznych, z. 300.
- Gołata E. (2004), *Estymacja pośrednia aktywności ekonomicznej na potrzeby spisu opartego na rejestrach*, Pomiar Informacji w Gospodarce, Zeszyty Naukowe 149, Wyd. UE, Poznań.
- GUS (2013), Bezrobocie rejestrowane w powiatach w 2003 r. Bank Danych Lokalnych. Odczyt z 23.08.2013.
- Kałużyńska M., Smyk K., Wiśniewski J. (red.) (2009), *Pięć lat Polski w Unii Europejskiej*, Urząd Komitetu Integracji Europejskiej, Warszawa.
- Paradysz J. (2008), *Kryteria dobroci estymacji dla małych obszarów*, [w:] *Statystyka społeczna – dokonania, szanse, perspektywy*, red. K. Jakóbik, Biblioteka Wiadomości Statystycznych, tom 57, Główny Urząd Statystyczny, Warszawa.
- Paradysz J., Paradysz K. (2011), *Benchmarking w statystyce małych obszarów*, [w:] *Klasyfikacja i analiza danych – teoria i zastosowanie*, red. K. Jajuga, M. Walesiak, Prace Naukowe Uniwersytetu Ekonomicznego nr 176, Taksonomia 18, Wydawnictwo UE, Wrocław.
- Paradysz K. (2013), *Benchmarkowa analiza estymacji dla małych obszarów na lokalnych rynkach pracy*, [w:] *Klasyfikacja i analiza danych – teoria i zastosowanie*, red. K. Jajuga, M. Walesiak, Prace Naukowe Uniwersytetu Ekonomicznego nr 279, Taksonomia 21, Wydawnictwo UE, Wrocław.
- Rao J.N.K. (2003), *Small Area Estimation*, Wiley.
- Ryś-Jurek R. (2012), *Zastosowanie hierarchicznej klasyfikacji aglomeracyjnej do grupowania krajów Unii Europejskiej ze względu na strukturę i skalę produkcji gospodarstw rolnych*, „Przegląd Statystyczny”, numer specjalny 2.
- Smętkowski M., Jałowiecki B., Gorzelak G. (2009), *Obszary metropolitalne w Polsce – diagnoza i rekomendacje*, „Studia Regionalne i Lokalne” nr 1(35).

## THE HIERARCHICAL METHOD OF GROUPING POWIATS AS A BENCHMARK APPROACH IN THE ASSESSMENT OF UNEMPLOYMENT BY BAEL IN SELECTED TYPES OF SMALL AREAS

**Summary:** One type of evaluation of the goodness of the results of the small areas estimation are the substantive criteria [Paradysz 2008]. For the estimation of unemployment rates for counties, substantive criterion (benchmark) can be identified in advance as some regularity in the local labor markets: metropolitan agglomerations, border counties, tourist areas. Using hierarchical clustering method we have made substantive assessment of the quality of the poviats unemployment rate estimation, which was performed earlier in the CSO [Bracha et al. 2004]. The results only partially confirm our assumptions, which means the quality of the estimation for small areas was not good enough.

**Keywords:** small area estimation (SAE), benchmarking, BAEL, labor market.