

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 385

Taksonomia 25

**Klasyfikacja i analiza danych –
teoria i zastosowania**

Redaktorzy naukowi

Krzysztof Jajuga

Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2015

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Tytuł dofinansowany ze środków Narodowego Banku Polskiego
oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Informacje o naborze artykułów i zasadach recenzowania
znajdują się na stronie internetowej Wydawnictwa
www.pracnaukowe.ue.wroc.pl
www.wydawnictwo.ue.wroc.pl

Publikacja udostępniona na licencji Creative Commons
Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 3.0 Polska
(CC BY-NC-ND 3.0 PL)



© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2015

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
e-ISSN 2392-0041 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Zamówienia na opublikowane prace należy składać na adres:
Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
tel./fax 71 36 80 602; e-mail:econbook@ue.wroc.pl
www.ksiegarnia.ue.wroc.pl

Druk i oprawa: TOTEM

Spis treści

Wstęp.....	9
Tomasz Bartłomowicz: Segmentacja konsumentów na podstawie preferencji wyrażonych uzyskanych metodą Maximum Difference Scaling	11
Barbara Batóg, Jacek Batóg, Andrzej Niemiec, Wanda Skoczylas, Piotr Waśniewski: Zastosowanie metod klasyfikacyjnych w identyfikacji kluczowych indyktorów osiągnięć w zarządzaniu wynikami przedsiębiorstw	20
Iwona Bąk: Wykorzystanie statystycznej analizy danych w badaniach turystyki transgranicznej na obszarach chronionych.....	28
Beata Bieszk-Stolorz: Ocena stopnia deprecjacji kapitału ludzkiego z wykorzystaniem nieliniowych modeli regresji.....	37
Mariola Chrzanowska, Nina Drejerska: Małe i średnie przedsiębiorstwa w strefie podmiejskiej Warszawy – określenie znaczenia lokalizacji z wykorzystaniem drzew klasyfikacyjnych.....	45
Adam Depta: Próba modelowania strukturalnego jakości życia osób jękających się jako konstrukt ukrytego na podstawie kwestionariusza SF-36v2	53
Katarzyna Dębowska: Wielowymiarowa analiza kondycji finansowej przedsiębiorstw sektora e-usług	63
Krzysztof Dmytrów, Mariusz Doszyń: Taksonomiczna procedura wspomagania kompletacji produktów w magazynie	71
Mariusz Doszyń, Sebastian Gnat: Propozycja procedury taksonomiczno-ekonometrycznej w indywidualnej wycenie nieruchomości.....	81
Marta Dziechciarz-Duda, Anna Król: Zastosowanie analizy <i>unfolding</i> i regresji hedonicznej do oceny preferencji konsumentów	90
Katarzyna Frodyma: Współzależność między poziomem rozwoju gospodarczego a udziałem energii ze źródeł odnawialnych w końcowym zużyciu w krajach Unii Europejskiej.....	99
Hanna Gruchociak: Porównanie struktury lokalnych rynków pracy wyznaczonych przy wykorzystaniu różnych metod w Polsce w latach 2006 i 2011 .	111
Alicja Grześkowiak, Agnieszka Stanimir: Postrzeganie środowiska pracy przez starszą i młodszą generację pracowników	120
Marta Hozer-Koćmiel, Christian Lis: Klasyfikacja krajów nadbałtyckich ze względu na czas prac wykonywanych w gospodarstwie domowym	129
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel: Zegar cyklu koniunkturalnego państw UE i USA w latach 1995-2013 w świetle badań synchronizacji.....	138
Aleksandra Łuczak: Wykorzystanie rozszerzonej interwałowej metody TOPSIS do porządkowania liniowego obiektów	147

Aleksandra Łuczak, Feliks Wysocki: Zintegrowane podejście do ustalania współczynników wagowych dla cech w zagadnieniach porządkowania linowego obiektów	156
Małgorzata Markowska, Danuta Strahl: Wykorzystanie klasyfikacji dynamicznej do identyfikacji wrażliwości na kryzys ekonomiczny unijnych regionów szczebla NUTS 2.....	166
Aleksandra Matuszewska-Janica, Marta Hozer-Koćmiel: Struktura zatrudnienia oraz wynagrodzenia kobiet i mężczyzn a przedmiotowa struktura gospodarcza w państwach UE.....	178
Anna M. Olszewska: Zastosowanie analizy korespondencji do badania związku pomiędzy zarządzaniem jakością a innowacyjnością przedsiębiorstw	187
Małgorzata Podogrodzka: Metoda aglomeracyjna w ocenie przestrzennego zróżnicowania starości demograficznej w Polsce	195
Ewa Roszkowska, Tomasz Wachowicz: Ocena ofert negocjacyjnych spoza dopuszczalnej przestrzeni negocjacyjnej.....	201
Ewa Roszkowska, Tomasz Wachowicz: Zastosowanie metody <i>unfolding</i> do wspomagania procesu negocjacji	210
Małgorzata Rószkiewicz: Próba diagnozy uwarunkowań poziomu wskaźnika braku odpowiedzi w środowisku polskich gospodarstw domowych.....	219
Marcin Salamaga: Próba identyfikacji muzycznych profili melomanów z wykorzystaniem drzew klasyfikacyjnych i regresyjnych	229
Agnieszka Sompolska-Rzechuła: Określenie czynników wpływających na prawdopodobieństwo poprawy poziomu rozwoju społecznego z wykorzystaniem modelu logitowego	239
Iwona Staniec: Wykorzystanie analizy czynnikowej w identyfikacji konstruktywów ukrytych determinujących ryzyko współpracy.....	248
Agnieszka Stanimir: Skłonność do zagranicznej mobilności młodszych i starszych osób	257
Mirosława Sztemberg-Lewandowska: Problemy decyzyjne w funkcjonalnej analizie głównych składowych.....	267
Tomasz Szubert: Demograficzno-społeczne determinanty określające subiektywny status jednostki w polskim społeczeństwie	276
Piotr Tarka: Własności 5- i 7-stopniowej skali Likerta w kontekście normalizacji zmiennych metodą Kaufmana i Rousseeuwa	286
Joanna Trzęsiok: Nielklasyczne metody regresji a problem odporności	296
Katarzyna Wawrzyniak: Ocena podobieństwa wyników uporządkowania województw uzyskanych różnymi metodami porządkowania	305
Katarzyna Wójcik, Janusz Tuchowski: Wykorzystanie metody opartej na wzorcach w automatycznej analizie opinii konsumenckich.....	314
Anna Zamojska: Zastosowanie analizy falkowej w ocenie efektywności funduszy inwestycyjnych	325

Summaries

Tomasz Bartłomowicz: Segmentation of consumers based on revealed preferences obtained with the Maximum Difference Scaling method	19
Barbara Batóg, Jacek Batóg, Andrzej Niemiec, Wanda Skoczylas, Piotr Waśniewski: Application of classification methods to identify the key performance indicators of performance management	27
Iwona Bąk: The application of statistical data analysis in the studies of cross-border tourism in protected areas.....	36
Beata Bieszk-Stolorz: Evaluating human capital depreciation by means of non-linear regression models.....	44
Mariola Chrzanowska, Nina Drejerska: Small and medium enterprises in the Warsaw suburban zone – determination of a localization’s role using classification trees	52
Adam Depta: An attempt of structural modelling of the quality of life of stuttering people as a latent construct, based on SF-36v2 questionnaire ...	62
Katarzyna Dębowska: Multidimensional analysis of financial condition of e-business services	70
Krzysztof Dmytrów, Mariusz Doszyń: Taxonomic procedure of supporting order-picking of products in a warehouse	80
Mariusz Doszyń, Sebastian Gnat: Taxonomic and econometric methods in individual real estate evaluation.....	89
Marta Dziechciarz-Duda, Anna Król: The application of unfolding analysis and hedonic regression in the investigation of consumers’ preferences	98
Katarzyna Frodyma: Interdependence between the level of economic development and the share of renewable energy in gross final energy consumption in the European Union.....	110
Hanna Gruchociak: Comparison of local labour markets structure designated using different methods in Poland in 2006 and 2011 years.....	119
Alicja Grzeškowiak, Agnieszka Stanimir: Perception of working environment by older and younger generation of workers.....	128
Marta Hozer-Koćmiel, Christian Lis: Classification of the Baltic Sea Region countries due to the time of household work.....	137
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel: Business cycle clock for the EU and the USA in 1995-2013 in the light of synchronization research.....	146
Aleksandra Łuczak: The use of the extended interval TOPSIS methods for linear ordering of objects.....	155
Aleksandra Łuczak, Feliks Wysocki: Integrated approach for determining the weighting coefficients for features in issues of linear ordering of objects.....	165

Małgorzata Markowska, Danuta Strahl: The application of dynamic classification for the identification of vulnerability to economic crisis in the EU NUTS 2 regions	177
Aleksandra Matuszewska-Janica, Marta Hozer-Koćmiel: The structure of male and female employment and remuneration vs. the basic economy structure in the EU countries	186
Anna M. Olszewska: The application of the correspondence analysis for the study of the relations between quality management and innovation in the enterprises.....	194
Małgorzata Podogrodzka: Agglomeration method in the age and ageing in Poland by voivodships.....	200
Ewa Roszkowska, Tomasz Wachowicz: Scoring the negotiation offers from the outside of the feasible negotiation space	209
Ewa Roszkowska, Tomasz Wachowicz: Application of the unfolding analysis to negotiation support.....	218
Małgorzata Rószkiewicz: An attempt to diagnose the determinants of non-response rate in Polish households surveys	228
Marcin Salamaga: Attempt to identify music lovers profiles using classification and regression trees	238
Agnieszka Sompolska-Rzechuła: The definition of factors influencing the probability of improving the level of human development using the logit model.....	247
Iwona Staniec: The use of factor analysis to identify hidden constructs – determinants of the cooperation risk	256
Agnieszka Stanimir: Willingness to mobility abroad among younger and older persons	266
Mirosława Sztemberg-Lewandowska: Decision problems in functional principal components analysis.....	275
Tomasz Szubert: Socio-demographic factors determining subjective social status of an individual in Polish society	285
Piotr Tarka: Normalization methods of variables and measurement on 5 and 7 point Likert scale	295
Joanna Trzęsiok: Non-classical regression methods vs. robustness	304
Katarzyna Wawrzyniak: The evaluation of the similarity of the voivodships' orderings obtained by means of different methods.....	313
Katarzyna Wójcik, Janusz Tuchowski: Using pattern-based opinion mining.....	324
Anna Zamojska: Mutual funds performance measurement – wavelets analysis approach.....	333

Aleksandra Matuszewska-Janica

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: aleksandra_matuszewska@sggw.pl

Marta Hozer-Koćmiel

Uniwersytet Szczeciński

e-mail: mhk@wneiz.pl

STRUKTURA ZATRUDNIENIA ORAZ WYNAGRODZENIA KOBIECI I MĘŻCZYŹN A PRZEDMIOTOWA STRUKTURA GOSPODARCZA W PAŃSTWACH UE

Streszczenie: W prezentowanym badaniu poddano analizie relację między przedmiotową strukturą gospodarczą a strukturą zatrudnienia kobiet i mężczyzn oraz ich wynagrodzeniami. Wyznaczone wskaźniki dotyczące rynku pracy i struktury gospodarczej posłużyły do klasyfikacji państw UE. Wykorzystano w tym celu metodę *k*-średnich. W badaniu szczególną uwagę skupiono na krajach nadbałtyckich i ich umiejscowieniu na tle pozostałych państw UE. Analizę przeprowadzono przy wykorzystaniu danych z Badania Struktury Wynagrodzeń (SES).

Słowa kluczowe: przedmiotowa struktura gospodarcza, struktura zatrudnienia kobiet, luka płacowa, kraje nadbałtyckie, metoda *k*-średnich.

DOI: 10.15611/pn.2015.385.19

1. Wstęp

Ważnym czynnikiem, który wpływa na funkcjonowanie rynku pracy, jest otoczenie gospodarcze (por. np. [Kumor 2010]), a struktura zatrudnienia jest ściśle powiązana z przedmiotową (rodzajową) strukturą gospodarczą. Dyskusje prowadzone na temat różnic w płacach kobiet i mężczyzn wskazują, że istotną rolę w tej kwestii odgrywa koncentracja zatrudnienia w określonych sektorach gospodarczych. Obserwuje się zjawisko tzw. segregacji zawodowej (por. np. [Levanon i in. 2009; Blau i in. 2010; Blau i in. 2013]). Kobiety częściej wybierają pracę w edukacji oraz służbie zdrowia i opiece społecznej, czyli sektorach usługowych. Natomiast więk-

szy odsetek mężczyzn jest zatrudnianych w górnictwie, budownictwie czy sektorze energetycznym. Sektory sfeminizowane odznaczają się mniejszą produktywnością niż te, w których przeważnie zatrudnia się mężczyzn. W związku z tym powstały dwa pytania. Po pierwsze, czy można wyodrębnić jednorodne grupy państw różniące się od siebie m.in. stopniem feminizacji sektorów gospodarczych oraz strukturą dysproporcji w wynagrodzeniach kobiet i mężczyzn? Po drugie, czy istnieje relacja między strukturą gospodarczą a stopniem feminizacji branż oraz między strukturą gospodarczą a dysproporcjami w wynagrodzeniach kobiet i mężczyzn?

Pierwszym celem artykułu było sprawdzenie możliwości wyodrębnienia jednorodnych grup państw UE ze względu na analizowane grupy zmiennych. Drugim celem była ocena podobieństwa struktury zatrudnienia kobiet i mężczyzn, luki w płacach implikowanej płcią oraz przedmiotowej struktury gospodarczej w państwach Unii Europejskiej ze szczególnym uwzględnieniem krajów nadbałtyckich (Dania, Finlandia, Szwecja, Estonia, Litwa, Łotwa, Polska)¹. Kraje te stanowią makroregion, dla którego konstruowana jest osobna strategia (Strategia Unii Europejskiej dla regionu Morza Bałtyckiego) oraz programy wsparcia finansowego, np. Program Region Morza Bałtyckiego. Kraje Skandynawskie będące liderami gospodarczymi regionu są jednocześnie wzorcem na świecie w zakresie stosowania polityki prorównościowej (por. np. [Hozer-Koćmiel 2007]). Prezentowana analiza miała umożliwić weryfikację hipotezy, że dla państw UE istnieje związek między przedmiotową strukturą gospodarczą a strukturą zatrudnienia kobiet i mężczyzn oraz poziomem ich wynagrodzeń.

W badaniu wykorzystano cztery grupy wskaźników. Pierwsza z nich to udziały wartości dodanej brutto (*Gross Value Added* – GVA) z danej branży w wartości dodanej brutto z analizowanych sektorów. W podejściu od strony produkcji wartość dodaną brutto wykorzystuje się w kalkulacji PKB. Jest on (PKB) sumą wartości dodanej brutto poszczególnych sektorów powiększoną o podatki od produktów i pomniejszoną o dotacje do produktów. Drugim rodzajem wskaźników jest udział kobiet wśród pracowników zatrudnionych w poszczególnych branżach. Pozwala to na ocenę stopnia feminizacji branż. Trzeci i czwarty typ wskaźników dotyczą poziomów wynagrodzeń w poszczególnych sektorach. Są to relacja przeciętnego wynagrodzenia w branży do przeciętnego wynagrodzenia w kraju oraz wskaźnik *Gender Pay Gap*.

2. Analizowane wskaźniki, dane i metodyka badania

W części badania dotyczącej struktury przedmiotowej gospodarki wykorzystano wskaźnik mierzący udział wartości dodanej brutto z danej branży i kraju w wartości dodanej brutto z analizowanych sektorów danego kraju:

¹ W analizie nie uwzględniono Niemiec, ponieważ występują tam znaczne różnice w analizowanych wskaźnikach między częściami wschodnią a zachodnią.

$$SGVA_{ij} = \frac{GVA_{ij}}{GVA_i} \cdot 100, \quad (1)$$

gdzie: $SGVA_{ij}$ – udział wartości dodanej brutto z j -tej branży oraz i -tego kraju w wartości dodanej brutto z analizowanych sektorów danego kraju; GVA_{ij} – wartość dodana brutto dla j -tej branży oraz i -tego kraju; GVA_i – wartość dodana brutto w i -tym kraju we wszystkich analizowanych branżach: $GVA_i = \sum_{j=C}^O GVA_{ij}$.

Indeks j oznacza sektor gospodarczy: $j = C$ (górnictwo), D (przetwórstwo przemysłowe), E (energetyka), F (budownictwo), G (handel i naprawy), H (hotele i restauracje), I (logistyka), J (pośrednictwo finansowe), K (obsługa nieruchomości i firm), M (edukacja), N (ochrona zdrowia i pomoc społeczna), O (działalność usługowa, pozostała)².

Indeks i oznacza kraj: $i = AT$ (Austria), BE (Belgia), DE (Niemcy), DK (Dania), ES (Hiszpania), FI (Finlandia), FR (Francja), GR (Grecja), IE (Irlandia), IT (Włochy), LU (Luksemburg), NL (Holandia), PT (Portugalia), SE (Szwecja), BG (Bułgaria), CY (Cypr), CZ (Czechy), EE (Estonia), HU (Węgry), LT (Litwa), LV (Łotwa), MT (Malta), PL (Polska), RO (Rumunia), SI (Słowenia), SK (Słowacja)³.

Jako miarę struktury zatrudnienia kobiet przyjęto udział kobiet wśród pracowników zatrudnionych w poszczególnych branżach i krajach:

$$SE_{F,ij} = \frac{E_{F,ij}}{E_{F,ij} + E_{M,ij}} \cdot 100, \quad (2)$$

gdzie: $SE_{F,ij}$ – udział kobiet w zatrudnieniu w grupie pracowników z i -tego kraju pracowników i j -tej branży; $E_{F,ij}$ – liczba zatrudnionych kobiet w i -tym kraju i j -tej branży; $E_{M,ij}$ – liczba zatrudnionych mężczyzn w i -tym kraju i j -tej branży; indeksy i i j oznaczają to samo co we wzorze (1).

Do pomiaru poziomu wynagrodzeń w poszczególnych branżach wykorzystano relację przeciętnego wynagrodzenia brutto za godzinę w i -tym kraju i j -tej branży do przeciętnego wynagrodzenia brutto za godzinę w i -tym kraju:

$$SGHE_{ij} = \frac{GHE_{ij}}{GHE_i} \cdot 100, \quad (3)$$

gdzie: GHE_{ij} – przeciętne wynagrodzenie brutto za godzinę w i -tym kraju i j -tej branży; GHE_i – przeciętne wynagrodzenie brutto za godzinę w i -tym kraju; indeksy i i j oznaczają to samo co we wzorze (1).

² W prezentowanej analizie przyjęto klasyfikację sektorów gospodarczych według NACE rev. 1.1.

³ Ze względu na brak danych w bazach Eurostatu nt. GVA w 2006 r. z analizy wykluczono Wielką Brytanię.

Jako miernik różnic w płacach kobiet i mężczyzn przyjęto wskaźnik *Gender Pay Gap* wyznaczany zgodnie ze wzorem (por. np. [Witkowska 2013]):

$$GPG_{ij} = \frac{GHE_{M,ij} - GHE_{F,ij}}{GHE_{M,ij}} \cdot 100, \quad (4)$$

gdzie: $GHE_{M,ij}$ – przeciętne wynagrodzenie mężczyzn brutto za godzinę w i -tym kraju i j -tej branży; $GHE_{F,ij}$ – przeciętne wynagrodzenie kobiet brutto za godzinę w i -tym kraju i j -tej branży; indeksy i i j oznaczają to samo co we wzorze (1).

Wybór wskaźników (2)-(4) bazuje na wcześniej przeprowadzonych analizach, w których zidentyfikowano zależność między dysproporcjami w płacach kobiet i mężczyzn a stopniem feminizacji grup zawodowych oraz poziomem wynagrodzeń (por. np. [Matuszewska-Janica 2014]).

Dane dotyczące wynagrodzeń i zatrudnienia zaczerpnięto z Badania Struktury Wynagrodzeń (*Structure of Earnings Survey* – SES) przeprowadzonego w 2006 roku. Są to dane w wersji zagregowanej (tzw. *metadata*) udostępnione przez Eurostat. SES jest badaniem przeprowadzanym co 4 lata (począwszy od 2002 roku), gdzie respondentami są firmy zatrudniające co najmniej 10 osób. Zaletą tego badania jest to, że pozyskane informacje pochodzą z rejestrów tych przedsiębiorstw. Dane o wartości dodanej brutto (GVA) również pochodzą z baz danych Eurostatu.

Prezentowane wyniki dotyczą grupowania 26 państw UE (bez Wielkiej Brytanii) ze względu na wielkość wskaźników (1)-(4) w poszczególnych branżach (dla każdego typu wskaźnika przeprowadzono osobne grupowanie). Wykorzystano w tym celu jedną z najczęściej stosowanych metod analizy skupień – metodę k -średnich (por. [McQueen 1967; Gatnar, Walesiak (red.) 2004]), zaimplementowaną w programie STATISTICA (z opcją maksymalizacji odległości skupień). Metodę postępowania przy przeprowadzaniu analizy skupień zaczerpnięto z pracy [Walesiak 2006]. Do normalizacji zmiennych wykorzystano standaryzację, a do wyznaczenia odległości – odległość euklidesową.

Dla każdego rodzaju zmiennej (opisanych wzorami (1)-(4)) przeprowadzono grupowania, w których obiekty były klasyfikowane przy z góry ustalonej liczbie skupień wynoszącej od 2 do 10 ($k = 2, 3, \dots, 10$). Za referencyjną liczbę skupień przyjęto 7. Dobór ten bazował na koncepcji podziału krajów europejskich przedstawionej w pracy [Ostergren i Le Bossé 2011, s. 21-36]. Na jej podstawie dla krajów UE można wydzielić siedem regionów, które wynikają z uwarunkowań geograficznych, historycznych i kulturowych: (1) Europę Północną (Dania, Szwecja, Finlandia, Estonia i Łotwa), (2) Wielką Brytanię i Irlandię, (3) Europę Zachodnią (Francja, Holandia, Belgia i Luksemburg), (4) Europę Środkowo-Zachodnią (Niemcy, Austria), (5) Europę Środkowo-Wschodnią (Polska, Litwa, Czechy, Słowacja, Słowenia i Węgry), (6) Bałkany (Bułgaria, Rumunia, Grecja i Cypr) oraz (7) Europę Zachodnią basenu Morza Śródziemnego (Portugalia, Hiszpania, Włochy,

Malta). W tej koncepcji rozważane w analizie kraje nadbałtyckie zostały przydzielone do dwóch regionów: Europy Północnej i Europy Środkowo-Wschodniej.

Dla każdej z otrzymanych klasyfikacji wyznaczono takie charakterystyki, jak indeks Calińskiego i Harabasa (C-H, por. [Caliński, Harabasz 1974]), indeks Huberta i Levina (H-L, por. [Hubert, Levin 1976]), oraz wskaźnik Rousseeuwa (SI, tzw. *silhouette index*, por. [Kaufman, Rousseeuw 1990]). Ocenę podobieństwa klasyfikacji państw UE przeprowadzono przy wykorzystaniu skorygowanej miary Randa (ARI, por. [Hubert, Arabie 1985]).

3. Wyniki

W tabeli 1 zaprezentowano dla wszystkich grup zmiennych (*SVGA*, *GPG*, *SE*, *SGHE*) wartości wskaźników SI dla referencyjnej liczby skupień (równiej 7) oraz liczbę skupień, dla których uzyskano najlepsze wartości indeksów Calińskiego i Harabasa (C-H), Huberta i Levine (H-L) oraz Rousseeuwa (SI). Wskazania poszczególnych miar znacznie odbiegają od siebie. Może to świadczyć o problemach w wyodrębnieniu sferycznych grup (przy zastosowaniu metody *k*-średnich). Na słabą strukturę klas wskazują niskie wartości indeksów SI (nie przekraczają 0,4). Ponadto dla każdej grupy zmiennych wskaźnik SI wskazywał inną liczbę skupień jako najlepszą. Największą wartość wskaźnika Rousseeuwa uzyskano dla podziału państw UE na dwie grupy na podstawie wartości *GVA* w poszczególnych branżach (SI = 0,39). W przypadku podziału ze względu na strukturę luki płacowej w sektorach najlepszym podziałem okazał się ten na siedem grup (SI = 0,26). Z kolei za najlepszy podział ze względu na strukturę feminizacji branż można uznać klasyfikację do trzech grup (SI = 0,30). W klasyfikacji ze względu na wysokość przeciętnego wynagrodzenia w branżach podziałem z największą wartością wskaźnika SI (SI = 0,24) okazał się ten na 9 grup. Dla referencyjnej liczby skupień jedynie w przypadku grupowania ze względu na poziom wskaźnika *GPG* w poszczególnych branżach zdiagnozowano słabą strukturę klas (SI = 0,26). Dla pozostałych grup zmiennych taka struktura klas nie występowała (SI były mniejsze od 0,25).

Powodem słabej struktury klas może być to, że struktura *GVA*, stopnia feminizacji, przeciętnych wynagrodzeń i nierówności w płacach kobiet i mężczyzn w poszczególnych sektorach jest specyficzna dla każdego państwa. W związku z tym sytuację kobiet na rynku pracy należałoby rozpatrywać osobno w odniesieniu do każdego kraju lub sektora.

W przypadku dwóch grup zmiennych: uwzględniających (1) strukturę gospodarczą (*SVGA*) oraz (2) stopień feminizacji sektorów (*SE*) dla uzyskanych podziałów rysuje się wyraźna różnica między krajami „starej piętnastki” i byłymi krajami socjalistycznymi (przyjmowanymi do struktur UE od 2004 roku). Dla dwóch pozostałych grup zmiennych, tj. uwzględniających (3) poziom luki płacowej (*GPG*)

oraz (4) poziom wynagrodzeń (*SGHE*) w poszczególnych sektorach, taki podział nie jest widoczny. W odniesieniu do państw BSR nie zdiagnozowano wyraźnych podobieństw między krajami tego regionu. Zauważono jednak, że Dania razem ze Szwecją oraz Litwa z Łotwą należą do tych samych skupień w większości klasyfikacji.

Faktem wartym uwagi jest to, że w krajach lepiej rozwiniętych gospodarczo obserwuje się znacznie wyższy udział w wartości dodanej brutto sektora zdrowotnego (w szczególności w krajach skandynawskich). Z kolei w krajach tych wynagrodzenia w tym sektorze są przeciętnie niższe od średniej krajowej niż w krajach słabiej rozwiniętych.

Tabela 1. Liczba skupień, dla których uzyskano najlepsze wartości indeksów C-H, H-L i SI oraz wartości SI dla 7 skupień

Grupa zmiennych	Wartość indeksu SI dla $k = 7$	Liczba skupień			
		C-H	H-L	SI	wartość indeksu SI
<i>SGVA</i>	0,14	4	5	2	0,39
<i>GPG</i>	0,26	2	3	7	0,26
<i>SE</i>	0,11	2	2	3	0,30
<i>SGHE</i>	0,20	2	10	9	0,24

Źródło: opracowanie własne.

Tabela 2. Wartości skorygowanej miary Randa (ARI)

Grupy zmiennych		Liczba skupień								
		2	3	4	5	6	7	8	9	10
<i>GPG</i>	<i>SE</i>	-0,034	-0,025	-0,021	0,013	0,035	0,061	0,091 ^a	0,081	0,036
<i>GPG</i>	<i>SGHE</i>	-0,033	0,006	0,097	0,071	0,166 ^b	0,113	0,087	0,182 ^b	0,211 ^a
<i>GPG</i>	<i>SGVA</i>	-0,021	0,005	-0,013	0,055	0,115 ^a	0,098	0,088	0,094	0,090
<i>SE</i>	<i>SGHE</i>	-0,031	0,037	-0,046	-0,070	-0,011	0,013	0,107 ^a	0,055	0,039
<i>SE</i>	<i>SGVA</i>	0,000	0,250	0,337 ^{a,b}	0,293 ^b	0,145	0,160 ^b	0,186 ^b	0,121	0,300 ^b
<i>SGHE</i>	<i>SGVA</i>	0,068 ^b	0,263 ^{a,b}	0,004	-0,012	0,008	-0,013	0,074	0,004	0,020

Źródło: opracowanie własne. ^a – najwyższa wartość wskaźnika ARI dla danej pary grup zmiennych; ^b – najwyższa wartość wskaźnika ARI dla danej liczby skupień.

Do oceny podobieństwa między uzyskanymi podziałami państw wykorzystano skorygowaną miarę Randa, której wartości zostały zaprezentowane w tab. 2. Najwyższą wartość odnotowano dla relacji między klasyfikacją do 4 grup ze względu na poziom feminizacji oraz strukturę gospodarczą (ARI = 0,337). Dla tej pary grup zmiennych w większości grupowań odnotowano największą wartość ARI. Można to interpretować tak, że w krajach UE ze strukturą gospodarczą najbardziej powiązany jest stopień feminizacji branż. Z kolei dla pary grup zmiennych *GPG* i *SE* oraz *SE* i *SGHE* wartości wskaźników ARI nie przekroczyły 0,11. Świadczy to

o słabym powiązaniu klasyfikacji ze względu na stopień feminizacji i poziom luki płacowej oraz ze względu na stopień feminizacji i wysokość płac w sektorach. W przypadku grupowania ze względu na wartości wskaźnik *GPG* w branżach najsilniejsze powiązanie stwierdzono z klasyfikacją ze względu na poziom wynagrodzeń w sektorach.

Tabela 3. Wartości skorygowanej miary Randa (ARI) dla powiązań między klasyfikacją państw EU według Ostergrena i Le Bosségo a klasyfikacją otrzymaną metodą *k*-średnich

Grupa zmiennych	ARI
<i>GPG</i>	0,1152
<i>SE</i>	0,1219
<i>SGHE</i>	-0,0062
<i>SGVA</i>	0,0366

Źródło: opracowanie własne.

Wartości wskaźnika ARI obliczone między klasyfikacją państw Unii Europejskiej według Ostergrena i Le Bosségo a klasyfikacją otrzymaną metodą *k*-średnich ze względu na analizowane grupy zmiennych przedstawiono w tab. 3. Najwyższy poziom skorygowanej miary Randa otrzymano dla klasyfikacji ze względu na wartość luki płacowej (*GPG*) i ze względu na poziom feminizacji (*SE*). Jednakże relacje te są bardzo słabe – wartości miary ARI nie przekraczają 0,13. Nie stwierdzono powiązania między klasyfikacją wzorcową a klasyfikacjami ze względu na strukturę gospodarczą i strukturę wysokości wynagrodzeń (współczynniki ARI były bliskie zeru).

4. Podsumowanie

Otrzymane wyniki wskazują, że na podstawie badanych grup zmiennych nie można wyróżnić skupień państw o mocnej strukturze. Może to świadczyć o dużej specyficzności analizowanych zmiennych dla każdego rozważanego państwa. Grupowania o silniejszej strukturze klas otrzymano dla klasyfikacji ze względu na strukturę gospodarczą (dla grupy zmiennych *GVA*) oraz ze względu na stopień feminizacji poszczególnych sektorów gospodarczych (dla grupy zmiennych *SE*). Oznacza to, że państwa UE w większym stopniu charakteryzują się podobieństwem w odniesieniu do struktury gospodarczej oraz zatrudnienia kobiet w branżach niż pod względem poziomu wynagrodzeń w branżach i ich zróżnicowania ze względu na płęć. Taki stan rzeczy może też rzutować na wyniki dotyczące analizy podobieństwa uzyskanych klasyfikacji. Skorygowana miara Randa wskazała, że najbardziej podobne klasyfikacje otrzymano dla grup zmiennych odnoszących się do struktury gospodarczej (*SVGA*) i stopnia feminizacji sektorów (*SE*). Potwierdza to, że istnieją relacje między przedmiotową strukturą gospodarki i sytuacją kobiet na rynku

pracy związaną ze strukturą zatrudnienia kobiet i mężczyzn (stopniem feminizacji sektorów).

Otrzymane wyniki wskazują na duże zróżnicowanie sytuacji kobiet i mężczyzn na rynku pracy zarówno w krajach nadbałtyckich, jak i w pozostałych krajach UE. Można jednak dostrzec dość wyraźny podział między analizowanymi wskaźnikami dla dwóch grup: krajów tzw. starej piętnastki oraz państw dawnego bloku postkomunistycznego. W przypadku krajów nadbałtyckich należy zwrócić uwagę na fakt, że w prezentowanych klasyfikacjach do tych samych skupień były przyporządkowywane Dania razem ze Szwecją oraz Litwa z Łotwą, co świadczy o podobieństwie tych krajów ze względu na analizowane zmienne.

Na podstawie przedstawionych wyników badań planowane są dalsze analizy porównawcze struktury gospodarczej i sytuacji kobiet i mężczyzn na rynku pracy. Będą one bazować na bardziej szczegółowych charakterystykach zatrudnienia i wynagrodzenia oraz na danych o mniejszym stopniu agregacji.

Literatura

- Blau F.D., Brummund P., Liu A.Y., 2013, *Trends in occupational segregation by gender 1970–2009: adjusting for the impact of changes in the occupational coding system*, Demography, vol. 50, no. 2, s. 471-492.
- Blau F.D., Ferber M.A., Winkler A.E., 2010, *The Economics of Women, Men, and Work*, Upper Saddle River, NJ: Prentice Hall.
- Calinski R.B., Harabasz J., 1974, *A Dendrite method for cluster analysis*, Communications in Statistics, vol. 3, no. 1, s. 1-27.
- Gatnar E., Walesiak M. (red.), 2004, *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo AE we Wrocławiu, Wrocław.
- Hozer-Koćmiel M., 2007, *Gender Mainstreaming w ekonomii. Dystrybucja czasu i wartości pracy kobiet*, Instytut Analiz, Diagnoz i Prognoz Gospodarczych, Szczecin.
- Hubert L.I., Levine J.R., 1976, *Evaluating object set partitions: free sort analysis and some generalizations*, Journal of Verbal Learning and Verbal Behaviour, vol. 15, no. 4, s. 459-470.
- Hubert L.J., Arabie P., 1985, *Comparing partitions*, Journal of Classification, vol. 2, no. 1, s. 193-218.
- Kaufman L., Rousseeuw P.J., 1990, *Finding Groups in Data: an Introduction to Cluster Analysis*, Wiley, New York.
- Kumor P., 2010, *Zależność nierówności płac od poziomu rozwoju gospodarczego*, Gospodarka Narodowa, nr 7-8, s. 45-62.
- Levanon A., England P., Allison P., 2009, *Occupational feminization and pay: assessing causal dynamics using 1950–2000 U.S. Census Data*, Social Forces, vol. 88, s. 865-981.
- MacQueen J.B., 1967, *Some Methods for Classification and Analysis of Multivariate Observations*, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, vol. 1, s. 281-297.
- Matuszewska-Janica A., 2014, *Wages inequalities between men and women: Eurostat SES metadata analysis applying econometric models*, Quantitative Methods in Economics, vol. XV, no. 1, s. 113-124.

- Ostergren R.C., Le Bossé M., 2011, *The Europeans, Second Edition: A Geography of People, Culture, and Environment*, The Guilford Press, New York.
- Walesiak M., 2006, *Rekomendacje w zakresie strategii postępowania w procesie klasyfikacji zbioru obiektów*, XXVII seminarium nt. „Przestrzenno-czasowe modelowanie zjawisk gospodarczych”, s. 185-203.
- Witkowska D., 2013, *Gender Disparities in the Labor Market in the EU*, International Advances in Economic Research, vol. 19, no. 4, s. 331-354.

THE STRUCTURE OF MALE AND FEMALE EMPLOYMENT AND REMUNERATION VS. THE BASIC ECONOMY STRUCTURE IN THE EU COUNTRIES

Summary: This paper analyzes the relationship between the economic structure and the employment structure of men and women and their wages. The classification of EU countries was accomplished using the labor market and the economic structure indicators. For that purpose the *k*-means method was used. Special attention was given to the Baltic Sea Region countries and their position in relation to other EU states. The analysis was based on the data from the Structure of Earnings Survey (SES).

Keywords: economy structure, women's employment structure, wage gap, BSR countries, *k*-means method.