



Jarosław Wąsowicz

Uniwersytet Ekonomiczny w Katowicach
Wydział Ekonomii
Katedra Analiz i Prognozowania Rynku Pracy
jaroslaw.wasowicz@ue.katowice.pl

PODOBIENSTWA RYNKÓW PRACY W GRUPIE KRAJÓW UE-28

Streszczenie: W artykule uwagę skoncentrowano na grupowaniu rynków pracy w badanych krajach UE-28 w celu utworzenia jednorodnych klas. Procedura grupowania uwzględnia wewnętrzne zróżnicowanie rynków pracy. W badaniu zastosowano hierarchiczne metody aglomeracyjne i metodę k -średnich. Głównymi kryteriami decydującymi o przynależności rynków pracy krajów UE-28 do skupień są wskaźnik zatrudnienia oraz współczynnik aktywności zawodowej.

Słowa kluczowe: rynki pracy w UE-28, metody aglomeracji, metoda k -średnich.

Wprowadzenie

Gospodarki narodowe, wchodzące w skład Unii Europejskiej są podmiotami badań o charakterze porównawczym. Badania tego typu mogą się sprowadzać do identyfikacji zróżnicowania pomiędzy krajami lub identyfikacji istniejących podobieństw, wraz z formalnym zakwalifikowaniem poszczególnych krajów do określonych skupień. Zarówno w pierwszym, jak i drugim przypadku, operacjonalizacja badań uwzględnia szereg cech diagnostycznych, charakteryzujących badane obiekty.

Zakres przedmiotowy artykułu obejmuje wybrane zmienne, charakteryzujące rynki pracy 28 krajów członkowskich Unii Europejskiej. Na podstawie kryteriów merytorycznego i formalno-statystycznego do zbioru cech diagnostycznych w badaniu zakwalifikowano pięć zmiennych, reprezentujących podażową i popytową sytuację na rynkach pracy krajów UE-28 w 2013 r. Obok kryterium merytorycznego, kryteriami rozstrzygającymi o kwalifikacji zmiennych diagnostycznych do badania były dostępność, porównywalność i aktualność danych

statystycznych. Źródłem danych była baza Eurostatu. Najbardziej aktualne dane (najczęściej publikowane przez Eurostat z dwuletnim opóźnieniem) za 2014 r. nie były dostępne w chwili rozpoczęcia badań. Badanie przeprowadzono z wykorzystaniem modułu „Analiza skupień” w pakiecie „Statistica”.

Celem artykułu jest identyfikacja statystycznie istotnych zmiennych, opisujących podobieństwa rynków pracy krajów UE-28, a także tych, w największym stopniu różnicujących badane obiekty. Operacjonalizacja tego celu sprowadza się do pogrupowania badanych rynków pracy krajów UE-28 w celu utworzenia jednorodnych klas, uwzględniających ich wewnętrzne zróżnicowanie przy wykorzystaniu hierarchicznych i niehierarchicznych metod wielowymiarowej analizy porównawczej (metod aglomeracyjnych: pełnego wiązania, Warda i metody k -średnich).

Przyпуска się, że zmienne opisujące popytową stronę rynków pracy objętych badaniem, w największym stopniu determinują istniejące pomiędzy nimi podobieństwa.

1. Grupowanie rynków pracy metodą pełnego wiązania i metodą Warda

Zgodnie z istniejącym dorobkiem naukowym w zakresie metod ilościowych, zbiorów zmiennych, charakteryzujących zjawisko, będących w szczególności agregatem ekonomicznym, podlega ocenie za pomocą kryteriów statystycznych [Panek, 2009, s. 18-19; Strahl, 1998, s. 45-59]. Zdolność dyskryminacyjną wszystkich zmiennych oceniono na podstawie wartości klasycznych współczynników zmienności.

$$|V(x_j)| \leq \varepsilon_{kr} \quad (1)$$

gdzie:

$\varepsilon_{kr} = 0,1$ i jest arbitralnie przyjętą wartością krytyczną klasycznego współczynnika zmienności.

W wyniku weryfikacji tak zdefiniowanej zdolności dyskryminacyjnej zmiennych, stwierdzono, że wszystkie zmienne charakteryzują się poziomem zmienności wyższym od przyjętego poziomu krytycznego.

W pierwszym etapie badań, tj. doborze zmiennych diagnostycznych, charakteryzujących podaźową i popytową sytuację na rynkach 28 państw członkowskich UE podjęto decyzję, że będzie ich sześć. Jednocześnie określono charakter tych zmiennych [Balcerowicz-Szkutnik, Sojka, 2011a, s. 10 i nast.; Mojena, 1997, s. 359 i nast.]:

X_1 – wskaźnik zatrudnienia osób w wieku 15-64 lat – stymulanta,

X_2 – współczynnik aktywności zawodowej ludności w wieku 15-64 lat (stymulanta),

X_3 – zatrudnieni zagrożeni ubóstwem (18-64 lat) – destymulanta,

X_4 – nowo zatrudnieni (udział osób posiadających zatrudnienie w okresie krótszym niż 12 miesięcy w ogólnej liczbie zatrudnionych) – stymulanta,

X_5 – stopa bezrobocia długoterminowego (15-74 lat) – destymulanta,

X_6 – bierni zawodowo (odsetek populacji, 15-64 lat) – destymulanta.

Oszacowane wartości miar zróżnicowania międzygrupowego oraz wewnątrzgrupowego zmiennych diagnostycznych oraz test F , weryfikujący statystyczną istotność pokazały, że współczynnik aktywności zawodowej i bierni zawodowo jako odsetek populacji są zmiennymi, dopełniającymi się do jedności (w obydwu przypadkach statystyka F przyjęła identyczną wartość na poziomie 9,91181). W rezultacie, zmienną X_6 wykluczono z badania empirycznego.

W celu transformacji destymulant w stymulanty wykorzystano przekształcenie różnicowe następującej postaci:

$$x_{ij}^S = a - bx_{ij}^D \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m; \quad b > 0 \quad (2)$$

gdzie:

$$b = 1$$

$$a = \max x_{ij}^D$$

Po przekształceniu wzoru (2) otrzymano:

$$x_{ij}^S = \max x_{ij} - x_{ij}^D \quad (3)$$

W celu pozbycia się miana zmiennych i ujednoczenia rzędu ich wielkości, przeprowadzono normalizację cech diagnostycznych, wykorzystując standaryzację stymulant. Procedurze tej poddano zbiór danych wyjściowych, stosując następującą formułę:

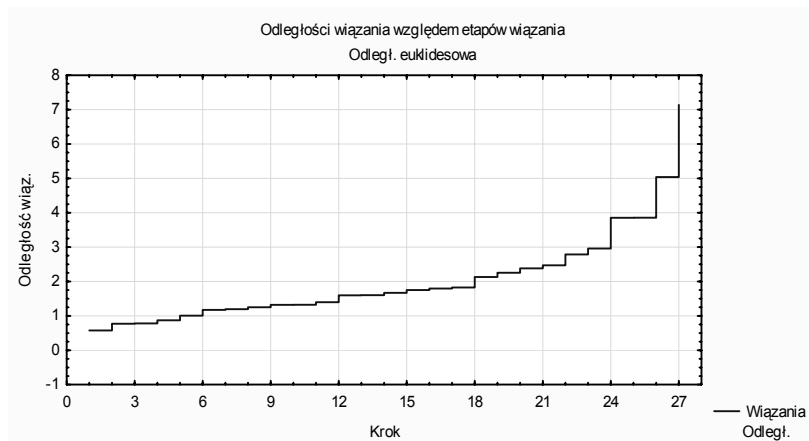
$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{S(x_j)} \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m \quad (4)$$

gdzie:

\bar{x}_j – średnia arytmetyczna j -tej zmiennej

$S(x_j)$ – odchylenie standardowe j -tej zmiennej.

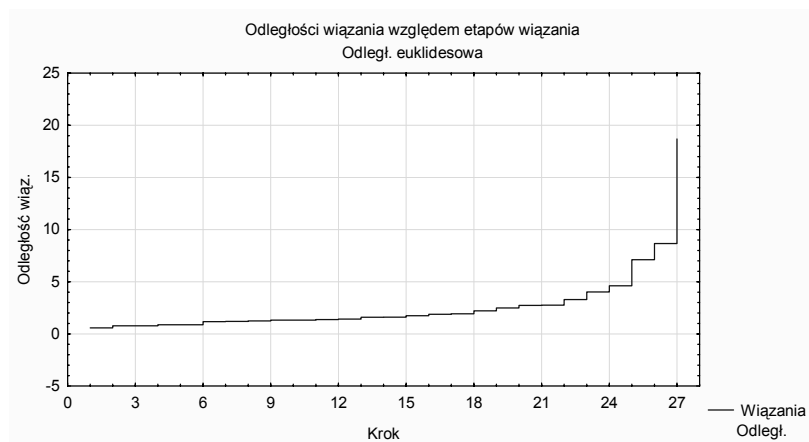
Oceny odległości pomiędzy skupieniami dokonano na podstawie liniowego wykresu, ilustrującego odległość euklidesową pomiędzy skupieniami państw UE-28, w momencie gdy były one łączone (rys. 1). Miejsce na wykresie przebiegu aglomeracji, w którym widoczne jest wyraźne spłaszczenie oznacza, że skupienia są odległe. W sensie statystycznym miejsce to jest najlepszym punktem tzw. odcięcia. W rezultacie, miejsce podziału dendrogramu należy umieścić po 27. kroku (najdłuższa linia pionowa).



Rys. 1. Wykres przebiegu aglomeracji (metoda pełnego wiązania).

Źródło: Na podstawie: [www 1].

W procesie grupowania państw członkowskich UE-28 wykorzystano także wykres przebiegu aglomeracji, uzyskany po dokonaniu porządkowania metodą Warda (rys. 2) [Balcerowicz-Szkutnik, Sojka, 2011b, s. 109-110]. Wyniki analizy tego wykresu są porównywalne do tych uzyskanych w metodzie pełnego wiązania. Miejsce podziału dendrogramu należy zlokalizować po 27. kroku (rys. 2).

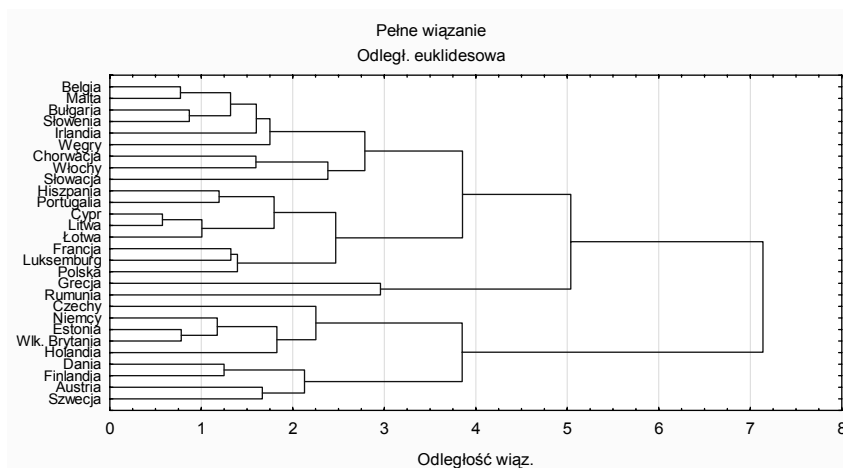


Rys. 2. Wykres przebiegu aglomeracji (metoda Warda)

Źródło: Na podstawie: [www 1].

Zastosowane metody hierarchiczne (pełnego wiązania, Warda) dały podstawę do podziału dendrogramów i sformułowania wniosku, że najlepszym rozwiązaniem jest przyjęcie dwóch grup państw UE-28 (rys. 3, 4). W przypadku metody pełnego wiązania utworzono podział na dwa następujące skupienia (rys. 3):

- skupienie 1: Belgia, Bułgaria, Chorwacja, Cypr, Francja, Grecja, Hiszpania, Irlandia, Litwa, Luksemburg, Łotwa, Malta, Polska, Portugalia, Rumunia, Słowacja, Słowenia, Węgry, Włochy,
- skupienie 2: Austria, Czechy, Dania, Estonia, Finlandia, Holandia, Niemcy, Szwecja, Wielka Brytania.



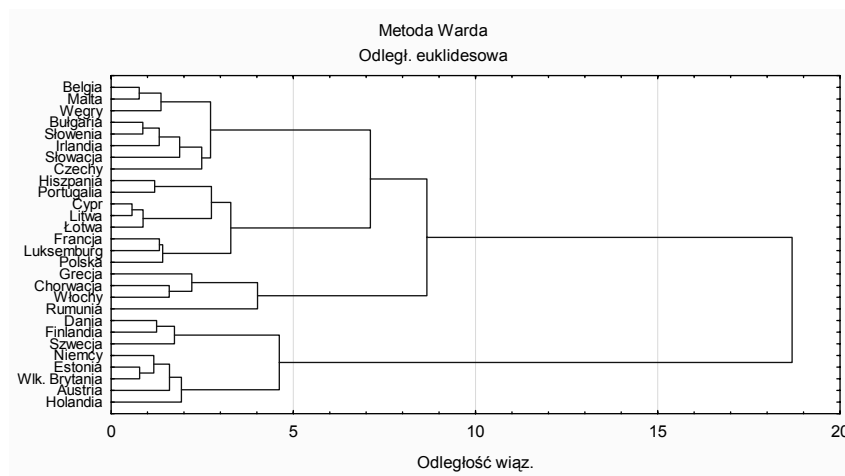
Rys. 3. Hierarchiczny wykres drzewa (metoda pełnego wiązania)

Źródło: Na podstawie: [www 1].

Wyniki uzyskane metodą Warda pozwoliły na wyłonienie dwóch grup badanych państw (rys. 4):

- grupa 1: Belgia, Bułgaria, Chorwacja, Cypr, Czechy, Francja, Grecja, Hiszpania, Irlandia, Litwa, Luksemburg, Łotwa, Malta, Polska, Portugalia, Rumunia, Słowacja, Słowenia, Węgry, Włochy,
- grupa 2: Austria, Dania, Estonia, Finlandia, Holandia, Niemcy, Szwecja, Wielka Brytania.

Różnica pomiędzy grupowaniem metodą najdalszego sąsiedztwa a metodą Warda sprowadza się do odmiennego zakwalifikowania gospodarki Czech. Zaskakujący jest fakt pojawienia się, takich państw jak Belgia, Francja, Irlandia i Luksemburg w jednym skupieniu z krajami Europy Środkowo-Wschodniej, np. Litwą, Łotwą, Polską, Słowacją, Węgrami oraz wieloma innymi państwami, znajdującymi się na poziomie rozwoju znacznie poniżej średniej unijnej, np. Chorwacją, Bułgarią, Rumunią. Wydaje się, że taka klasyfikacja badanych obiektów nie poddaje się sensownej interpretacji ekonomicznej.



Rys. 4. Hierarchiczny wykres drzewa (metoda Warda)

Źródło: Na podstawie: [www 1].

2. Grupowanie rynków pracy metodą k -średnich

W związku z koniecznością wstępnego określenia liczby skupień w metodzie k -średnich, wykorzystano wyniki uzyskane metodami podziału dendrogramów pełnego wiązania i Warda. W rezultacie, w pakiecie „Statistica” określono liczbę dwóch skupień i dziesięciu iteracji. Określona na wstępie liczba iteracji okazała się wystarczająca. Odległości pomiędzy państwami posortowano, a na początkowe centra skupień wybrano państwa przy stałych interwałach [Hartigan, Wong, 1979, s. 100-108]. Zastosowanie metody k -średnich pozwoliło na wyłonienie dwóch skupień rynków pracy w krajach UE-28 oraz ich odległości od środków skupień (tab. 1).

W tab. 1 widać, że w skupieniu drugim znalazły się dwa kraje, których odległości od środka całej grupy są relatywnie duże (Rumunia i Grecja). Wyniki wskazują, iż w grupie pierwszej występuje mniejsze zróżnicowanie odległości poszczególnych krajów od jej środka w porównaniu z grupą drugą. Wydaje się, że z punktu widzenia ekonomicznego kontekstu interpretacji wyników badania, metoda k -średnich charakteryzuje się większą wartością w porównaniu z wykorzystanymi wcześniej metodami hierarchicznymi (pełnego wiązania, Warda). Wynika to z mniejszego zróżnicowania wewnętrznego każdego z wyłonionych skupień rynków pracy krajów UE-28 i jednocześnie większego zróżnicowania międzygrupowego (tab. 1). W rezultacie, skład poszczególnych skupień nie budzi istotnych wątpliwości.

Tabela 1. Elementy skupień i odległości poszczególnych państw od środków skupień

Elementy skupienia numer 1 (14 przypadków)	Odległ.	Elementy skupienia numer 2 (14 przypadków)	Odległ.
Dania	0,991712	Bułgaria	0,270124
Niemcy	0,590950	Irlandia	0,743926
Estonia	0,317641	Grecja	1,050654
Francja	0,648051	Hiszpania	0,827460
Cypr	0,558516	Chorwacja	0,742585
Łotwa	0,624745	Włochy	0,584801
Litwa	0,581823	Węgry	0,663397
Luksemburg	0,890821	Malta	0,594314
Holandia	0,710153	Polska	0,528830
Austria	0,413992	Portugalia	0,582495
Finlandia	0,789486	Rumunia	1,532392
Szwecja	0,970806	Słowenia	0,483579
Wlk. Brytania	0,209144	Słowacja	0,769929

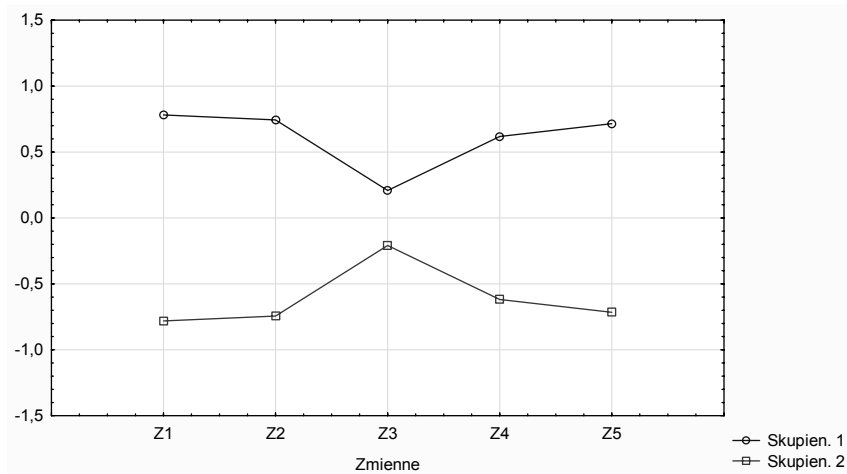
Źródło: Na podstawie: [www 1].

Tabela 2. Analiza wariancji zmiennych

Zmienna	Analiza wariancji					
	Między SS	df	Wewn. SS	df	F	Istotn. p
Z1	17,10522	1	9,89478	26	44,94653	0,000000
Z2	15,50301	1	11,49699	26	35,05948	0,000003
Z3	1,21989	1	25,78011	26	1,23030	0,277504
Z4	10,69518	1	16,30482	26	17,05475	0,000333
Z5	14,34674	1	12,65326	26	29,47979	0,000011

Źródło: Na podstawie: [www 1].

Weryfikacja wyników badania, uzyskanych metodą k -średnich wskazuje m.in., że zmienne diagnostyczne Z_1 , Z_2 , Z_4 i Z_5 są statystycznie istotne (przy poziomie istotności 0,01). Problem z poziomem statystycznej istotności występuje jedynie w przypadku zmiennej diagnostycznej Z_3 ($p = 0,277504$). Oszacowane wartości statystyki F oraz wartości p pozwalają stwierdzić, że głównymi kryteriami, różnicującymi badane rynki pracy krajów UE-28 i jednocześnie decydującymi o ich przynależności do skupień są zmienne Z_1 (wskaźnik zatrudnienia) oraz Z_2 (współczynnik aktywności zawodowej). Zmienna Z_3 (zatrudnieni zagrożeni ubóstwem) w największym stopniu przyczyniła się do zróżnicowania wewnątrzgrupowego w skupieniach, choć należy pamiętać o relatywnie wysokim prawdopodobieństwie popełnienia błędu obciążającego ten wniosek (relatywnie wysoka wartość p).



Rys. 5. Średnie wartości zmiennych dla poszczególnych grup państw

Źródło: Na podstawie: [www 1].

Elementem uzupełniającym proces weryfikacji uzyskanych wyników badania są wartości średnie obydwu skupień. Skupienie 1 charakteryzuje się wyższymi średnimi wartościami zestandaryzowanych zmiennych diagnostycznych. Największe różnice pomiędzy dwoma skupieniami rynków pracy w krajach UE-28 występują w przypadku zmiennych Z_1 (wskaźnik zatrudnienia) i Z_2 (współczynnik aktywności zawodowej). Oszacowane wartości średnie korespondują z uzyskanymi wynikami analizy wariancji międzygrupowych oraz wewnątrzgrupowych poszczególnych zmiennych diagnostycznych. Jednocześnie, najmniejsze różnice pomiędzy skupieniami rynków pracy występują w przypadku zmiennej Z_3 (zatrudnieni zagrożeni ubóstwem).

Podsumowanie

Wyniki badania uzyskane z zastosowaniem metod hierarchicznych i metody niehierarchicznej pozwoliły na sformułowanie wniosków, odnoszących się do problemu badawczego, jednocześnie zapewniając podstawę do weryfikacji postawionej we wprowadzeniu hipotezy badawczej:

1. Głównymi kryteriami decydującymi o przynależności rynków pracy krajów UE-28 do skupień są dwie zmienne: wskaźnik zatrudnienia oraz współczynnik aktywności zawodowej. Pierwsza z tych zmiennych jest zmienną popytową, zaś druga – zmienną opisującą podażową stronę badanych rynków pracy. W rezultacie, postawioną we wprowadzeniu hipotezę zweryfikowano jedynie częściowo, w odniesieniu do zmiennej popytowej.

2. Zmienną mającą najmniejszy wpływ na przynależność do grup jest odsetek osób zatrudnionych, zagrożonych ubóstwem.
3. Wyłoniona w wyniku zastosowania metody k -średnich grupa 1, charakteryzuje się wyższymi średnimi wartościami poszczególnych zmiennych diagnostycznych.
4. Wydaje się, że metoda k -średnich pozwoliła połączyć badane rynki pracy w jednorodne skupienia w sposób bardziej uzasadniony z punktu widzenia rzeczywistego, ekonomicznego ich zróżnicowania. Znacznie mniej adekwatne wyniki uzyskano za pomocą metod aglomeracyjnych (metody pełnego wiązania i metody Warda).

Literatura

- Balcerowicz-Szkutnik M., Sojka E. (2011a), *Pokolenie 50+ na europejskim rynku pracy – podsumowanie poszczególnych etapów badań*, Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach, Katowice.
- Balcerowicz-Szkutnik M., Sojka E. (2011b), *Metody ilościowe w finansach i rachunkowości. Wybrane zagadnienia*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.
- Hartigan J.A., Wong M.A. (1979), *A K-Means Clustering Algorithm*, „Journal of the Royal Statistical Society”, Series C (Applied Statistics), Vol. 28, No. 1.
- Mojena R. (1997), *Hierarchical Grouping Methods and Stopping Rules: An Evaluation*, „Computer Journal”, Vol. 20, No. 4.
- Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, Wydawnictwo Szkoły Głównej Handlowej, Warszawa.
- Strahl D. (red.) (1998), *Taksonomia struktur w badaniach regionalnych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- [www 1] <http://ec.europa.eu/eurostat> (dostęp: 22.06.2015).

THE SIMILARITY OF LABOUR MARKETS IN THE EU-28 COUNTRIES

Summary: The paper focuses on the clustering of labor markets in the EU-28 countries in order to form homogeneous clusters. The clustering procedure of EU countries is based on intra and inter-group differences in labor markets. The agglomerative hierarchical clustering and k -means clustering methods were used in the paper. The main criteria for deciding on countries belonging to clusters are employment rate and economic activity rate.

Keywords: labour markets in the EU-28, agglomerative hierarchical clustering method, k -means clustering.