

**Alicja Grześkowiak**

Uniwersytet Ekonomiczny we Wrocławiu

e-mail: alicja.grzeskowiak@ue.wroc.pl

---

**WYKORZYSTANIE WYBRANYCH FORM I TECHNIK  
WIZUALIZACJI DANYCH W BADANIACH STRUKTUR  
SPOŁECZNO-EKONOMICZNYCH**

---

**Streszczenie:** W artykule zaprezentowano wybrane metody wizualizacji, które mogą być przydatne podczas przeprowadzania analiz i porównywania struktur społeczno-ekonomicznych. Praca poświęcona jest technikom niewymagającym redukcji wielowymiarowości, której towarzyszy zazwyczaj utrata części informacji zawartych w zbiorze danych. Przedstawiono metody graficznej prezentacji struktur oraz ilustrowania ich podobieństwa i braku podobieństwa, podkreślając wady i zalety poszczególnych rozwiązań. Rozważania przeprowadzono na podstawie danych dotyczących struktury wykształcenia ludności krajów Unii Europejskiej.

**Słowa kluczowe:** analiza struktur społeczno-ekonomicznych, wizualizacja danych, struktura wykształcenia.

DOI: 10.15611/ekt.2014.2.03

## 1. Wstęp

Jednym z nurtów badań społeczno-ekonomicznych jest analiza struktur określonych zbiorowości. Struktura jest w tym przypadku traktowana jako „specyficzny obiekt złożony, scharakteryzowany nie wartościami różnych cech, lecz ciągiem wskaźników struktury  $p_1, p_2, \dots, p_r$ , takich, że  $\sum_{k=1}^r p_k = 1$ ” [Strahl (red.) 1998, s. 45]. Elementy  $p_k$  odzwierciedlają względne natężenie cechy w składnikach struktury, pozbawione są więc mian i sumują się do jedności, co pozwala na stosowanie swoistych metod pomiaru podobieństwa i odległości [Strahl (red.) 1998, s. 45]. Wymienione własności decydują również o preferencjach w zakresie graficznego sposobu przedstawiania struktur, które tradycyjnie ilustruje się za pomocą wykresów kołowych oraz skumulowanych wykresów słupkowych bądź kolumnowych. Tego rodzaju ilustracja bardzo dobrze sprawdza się w sytuacjach, w których należy zaprezentować strukturę jednego obiektu lub przeprowadzić analizę porównawczą dla kilku obiektów.

tów, natomiast staje się mniej czytelna, gdy przedmiotem analizy jest ich znaczna liczba. Ponieważ grafika, obok opisu słownego i tabel, stanowi bardzo użyteczny sposób przekazywania informacji statystycznej [Du Toit, Steyn, Stumpf 1986, s. 1], warto rozważyć inne techniki wizualizacji, które mogą mieć zastosowanie wspomagające badanie i porównywanie struktur.

Niniejsza praca realizuje dwa cele. Nadrzędnym celem poznawczym o charakterze metodologicznym jest przegląd alternatywnych sposobów graficznej prezentacji obiektów opisanych za pomocą ciągu wskaźników struktury oraz ilustrowania ich podobieństwa lub braku podobieństwa. Tak określone celowi zostały podporządkowane cele szczegółowe:

- wskazanie metod prezentacji elementów składowych struktur,
- wizualizacja podobieństwa struktur w powiązaniu z zastosowaniem metod grupowania,
- prezentacja graficzna braku podobieństwa struktur.

Realizowanym równoległym celem poznawczym o charakterze merytorycznym jest porównanie państw Unii Europejskiej pod względem struktury wykształcenia, będącej jedną z miar kapitału ludzkiego. Analizę przeprowadzono na podstawie danych Eurostatu<sup>1</sup>, zwracając szczególną uwagę na sytuację Polski na tle pozostałych krajów. Zgodnie z przyjętą przez Europejski Urząd Statystyczny metodologią definiuje się trzy zagregowane kategorie wykształcenia, odpowiadające określonym szczeblom edukacyjnym według Międzynarodowej Standardowej Klasyfikacji Wykształcenia (International Standard Classification of Education ISCED 1997), umożliwiającej porównywanie zagadnień z zakresu kształcenia w różnych krajach, a składającej się z siedmiu poziomów: ISCED 0-ISCED 6. Zestawienie agregatów przedstawiono w tab. 1.

**Tabela 1.** Rozpatrywane kategorie wykształcenia a poziomy klasyfikacji ISCED 1997

Nazwa kategorii	Poziomy
Niższe średnie	ISCED 0: edukacja przedszkolna, osoby bez wykształcenia ISCED 1: szkolnictwo podstawowe ISCED 2: szkolnictwo średnie I stopnia ISCED 3C: szkolnictwo średnie II stopnia o krótkim czasie nauki nieumożliwiające kontynuację edukacji na poziomie ISCED 5
Wyższe średnie	ISCED 3: szkolnictwo średnie II stopnia (z wyjątkiem ISCED 3C o krótkim czasie nauki) ISCED 4: szkolnictwo policealne niezaliczane do szkolnictwa wyższego
Wyższe	ISCED 5: szkolnictwo wyższe (pierwszy etap – profil akademicki i zawodowy) ISCED 6: szkolnictwo wyższe (drugi etap – zdobywanie kwalifikacji naukowych – stopień doktorski)

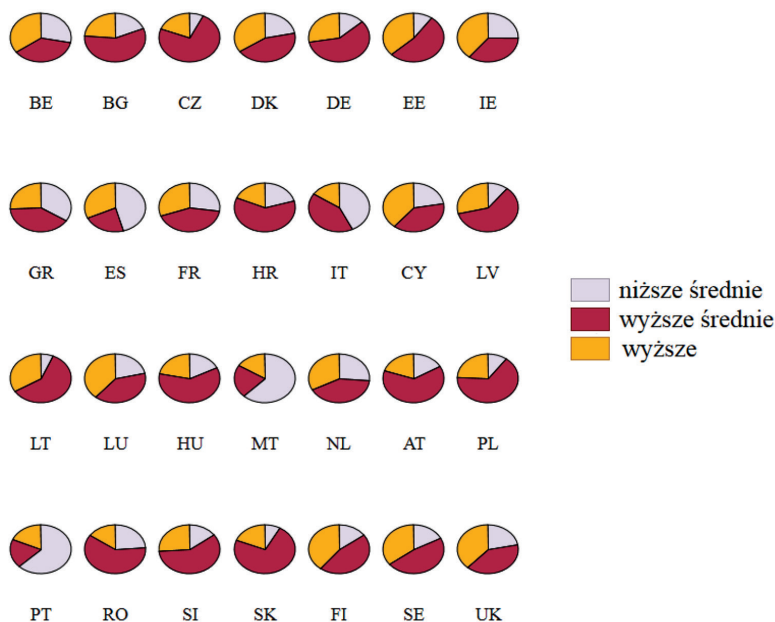
Źródło: opracowanie na podstawie [Dorośli w systemie edukacji formalnej... 2011, s. 80] oraz [[http://epp.eurostat.ec.europa.eu/cache/ITY\\_SDDS/en/edat\\_esms.htm](http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/edat_esms.htm)].

<sup>1</sup> Dane pochodzą z bazy [Educational attainment and outcomes...].

Należy zwrócić uwagę, że kategorię *niższe średnie* należy traktować jako „co najwyżej” ten rodzaj wykształcenia, gdyż zawiera ona osoby całkowicie bez wykształcenia oraz legitymujące się najniższymi jego poziomami, jak podstawowe oraz niższe średnie. Wykorzystane w pracy dane pochodzą z bazy Eurostatu stworzonej na podstawie LFS (*Labour Force Survey*) z 2012 r. i dotyczą osób w wieku 25-64 lat.

## 2. Wybrane formy i techniki graficznej prezentacji struktur

Wizualizację struktur bardzo często przeprowadza się za pomocą wykresów kołowych, skumulowanych kolumnowych oraz skumulowanych słupkowych. Jak nadmieniono we wstępie, tego rodzaju ujęcie jest odpowiednie i przejrzyste w przypadku ilustrowania jednej struktury bądź kilku struktur. W sytuacji porównywania większej liczby obiektów istnieje możliwość zastosowania zestawu wielu wykresów o tym samym charakterze. Przykładowo na rys. 1 przedstawiono prezentację gra-



AT – Austria, BE – Belgia, BG – Bułgaria, CZ – Czechy, CY – Cypr, DE – Niemcy, DK – Dania, EE – Estonia, ES – Hiszpania, FI – Finlandia, FR – Francja, GR – Grecja, HR – Chorwacja, HU – Węgry, IE – Irlandia, IT – Włochy, LT – Litwa, LU – Luksemburg, LV – Łotwa, MT – Malta, NL – Holandia, PL – Polska, PT – Portugalia, RO – Rumunia, SE – Szwecja, SI – Słowenia, SK – Słowacja, UK – Wielka Brytania.

**Rys. 1.** Ilustracja struktury ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej z zastosowaniem zestawu wykresów kołowych

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu STATISTICA ver. 10.

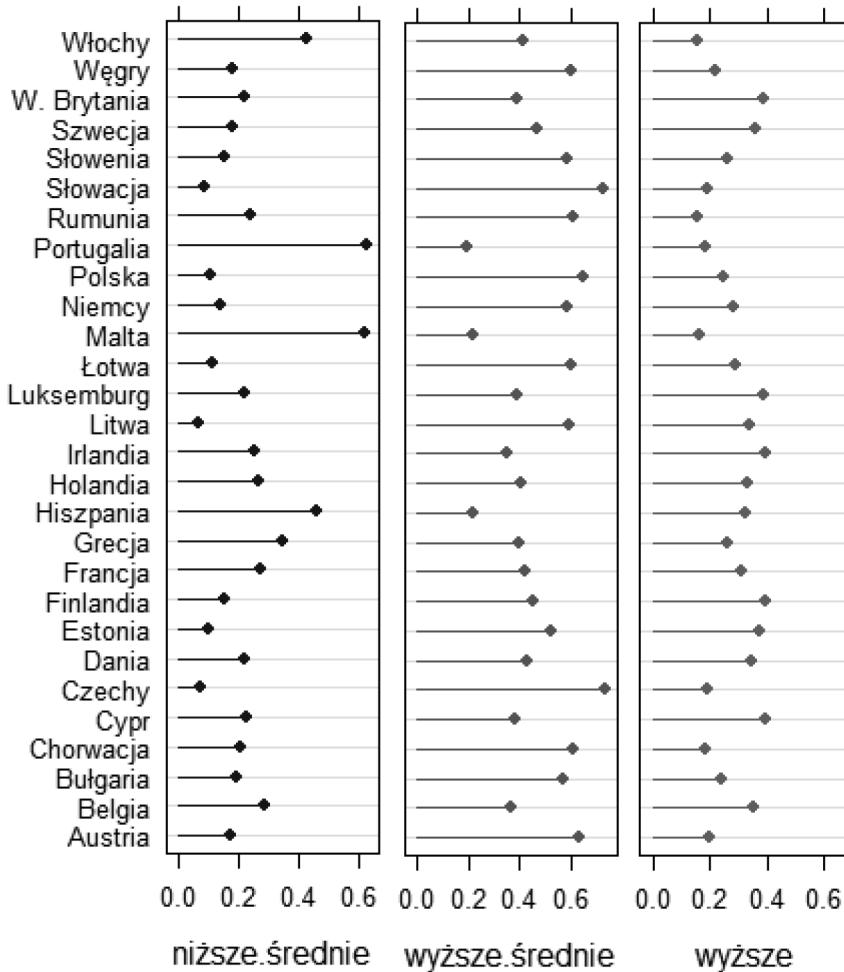
ficzną struktury wykształcenia w krajach Unii Europejskiej za pomocą tzw. wykresu obrazkowego składającego się z zestawu dwudziestu ośmiu wykresów kołowych. Analogicznie można przedstawić dane, używając zbioru wykresów kolumnowych.

Niewątpliwie tego rodzaju prezentacja wspomaga przekaz informacji, pozwalając uchwycić kluczowe prawidłowości, które nie byłyby wyraźnie widoczne w przedstawieniu tabelarycznym, ujmującym osiemdziesiąt cztery liczby. Przykładowo, korzystając z rys. 1., łatwo wskazać państwa, w których dominuje jeden z elementów struktury, np. bardzo wysoki odsetek osób z co najwyżej niższym średnim wykształceniem w Portugalii i na Malcie, czy znaczna przewaga osób z wykształceniem wyższym średnim w Czechach, Polsce i na Słowacji. Możliwe jest również odszukanie obiektów o najwyższych czy najniższych wartościach danego składnika, choć nie wynika to bezpośrednio z formy wykresu, lecz wymaga porównania ze sobą wszystkich elementów, co przy dużej ich liczbie byłoby bardzo utrudnione.

W celu porównania poszczególnych składowych struktur można się posłużyć trellisami, które stanowią „grupę narzędzi kompleksowej graficznej prezentacji (...) danych przestrzennych lub wielowymiarowych na jednym diagramie” [Suchecki (red.) 2010, s. 54]. Wizualizacja trellisowa przypomina swą konstrukcją prostokątną tabelę, zawierającą zestaw wykresów jednego typu, co umożliwia wyciągnięcie wniosków dotyczących zarówno wierszy, jak i kolumn [Suchecki (red.) 2010, s. 54], co w przypadku analizy struktur oznacza, że można przeprowadzać porównania wartości poszczególnych wskaźników struktury oraz obiektów. Bogate możliwości w zakresie tworzenia trellisów oferuje pakiet `lattice` dostępny w programie R, a ich obszerną charakterystykę można znaleźć w pracy Sarkara [2008]. Na rysunku 2 przedstawiono struktury będące przedmiotem rozważań za pomocą techniki trellisów z wykorzystaniem wykresów punktowych (*dotplot*), wzbogaconych o połączenia z punktem zerowym, co jest pomocne przy ocenie wartości relatywnych (por. [Sarkar 2008, s. 57-58]).

Ujęcie zaprezentowane na rys. 2 pozwala na szybką ocenę obiektów pod względem wartości wskaźników struktury oraz na wskazanie, jak duże są różnice występujące pomiędzy nimi. Zastosowanie jednakowej skali dla wszystkich wykresów umożliwia identyfikację elementów dominujących w poszczególnych obiektach. Rysunek 2 został sporządzony z zachowaniem porządku alfabetycznego, ale jeśli jest to pożądane ze względów interpretacyjnych, można obiekty uporządkować według wybranej kategorii. Układ punktów na rys. 2 unaocznia, że największe różnice dotyczą osób o najniższym poziomie wykształcenia. Bardzo niski odsetek ludności słabo wyedukowanej występuje na Litwie i w Czechach wobec wyraźnej przewagi tej części struktury w Portugalii i na Malcie. Najmniejsze dysproporcje odnotowano w przypadku osób o wyższym wykształceniu. Rozpatrując sytuację w Polsce, można stwierdzić, że na tle pozostałych państw odsetek osób o wykształceniu co najwyżej niższym średnim jest niewielki, o wyższym średnim – relatywnie duży, a o wyższym – przeciętny. Wydaje się, że wykresy typu *trellis* nadają się do badania pozycji jednego obiektu względem pozostałych, natomiast nie odzwierciedlają w bezpośredni

i przejrzysty sposób relacji pomiędzy wszystkimi obiektami ze względu na wszystkie wymiary.

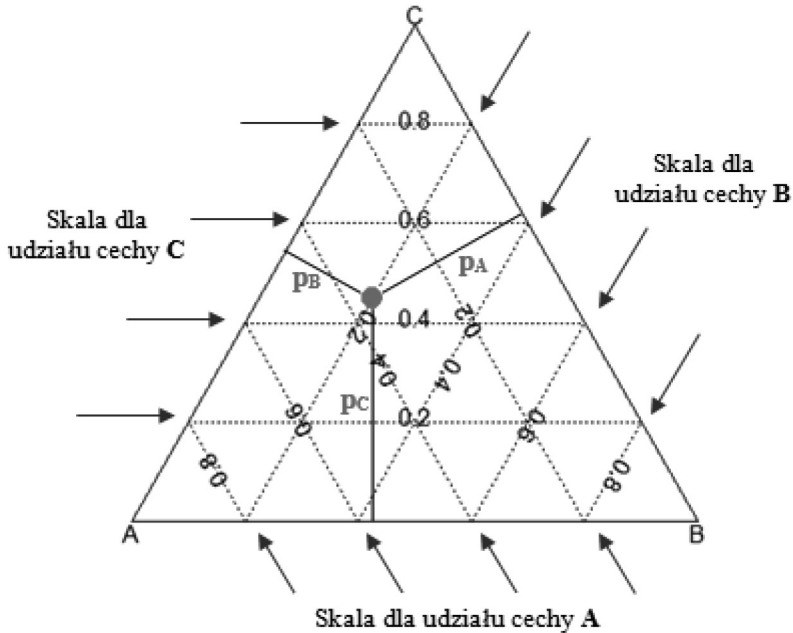


**Rys. 2.** Ilustracja struktury ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej z zastosowaniem zestawu wykresów typu *trellis*

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu R.

Alternatywnym rozwiązaniem w przypadku badania struktur trójelementowych mogą być wykresy trójkątne (zob. np. [Aitchison 1986, s. 5-7; Friendly 2000, s. 97-98]), bazujące na fakcie, że dla każdego wektora składającego się z trzech nieujemnych elementów  $(p_A, p_B, p_C)$ , spełniających warunek  $p_A + p_B + p_C = 1$ , istnieje dokładnie jeden punkt należący do trójkąta równobocznego o wysokości 1 odległy od jego boków o  $p_A, p_B, p_C$ , co daje jednoznaczne przedstawienie struktur trójele-

mentowych [Aitchison 1986, s. 6]. Idea tego rodzaju wizualizacji została zaprezentowana na rys. 3.

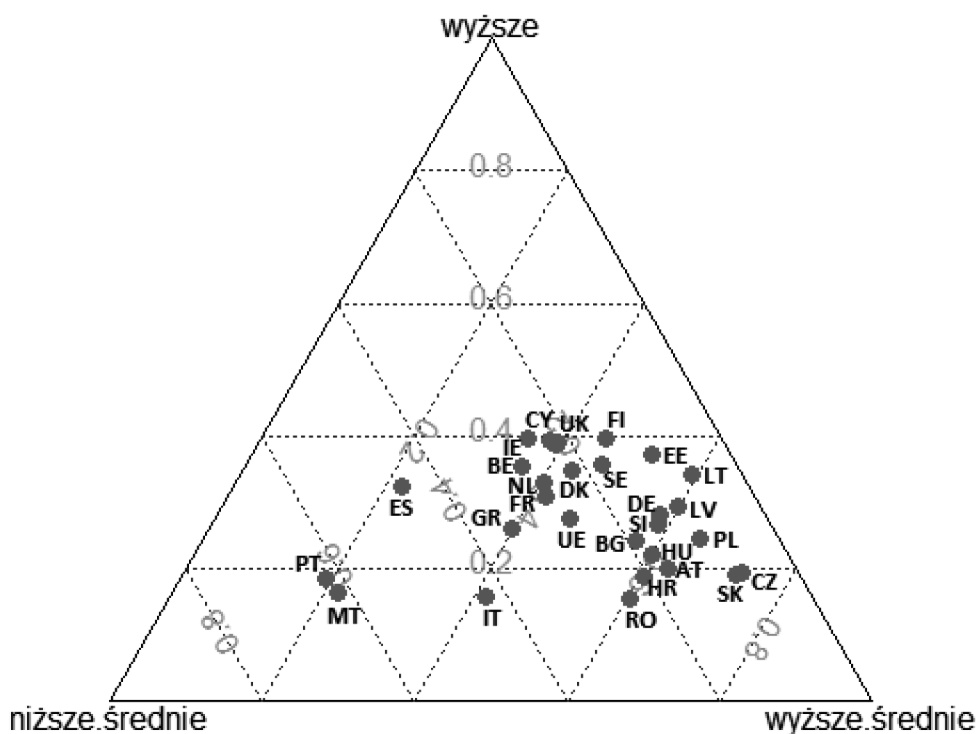


**Rys. 3.** Zasada konstrukcji wykresu trójkątnego odzwierciedlającego strukturę trójelementową  
Źródło: opracowanie własne z wykorzystaniem programu R.

Obiekty reprezentowane na wykresie trójkątnym mają trzy współrzędne, które mogą być odczytane na podstawie siatki znajdującej wewnątrz trójkąta lub skali zamieszczonej na jego bokach. Im bliżej danego wierzchołka znajduje się punkt, tym udział cechy przypisanej temu wierzchołkowi jest większy. Na rysunku 3 zaznaczono kółkiem przykładowy punkt, którego udziały wynoszą odpowiednio dla A: 0,35, dla B 0,2, dla C 0,45.

Odwzorowanie analizowanych struktur odnoszących się do wykształcenia ludności krajów UE za pomocą tego typu wykresu przedstawia rys. 4.

Wykres trójkątny umożliwi przegląd konfiguracji wielu obiektów ze względu na ich struktury. Pozwala zidentyfikować grupy obiektów podobnych ze względu na trzy cechy łącznie oraz elementy odstające. Naniesienie siatki z wartościami daje możliwość odczytania (przybliżonych) wartości udziałów i zorientowania się, które obiekty odznaczają się podobnym poziomem danego wskaźnika struktury. Układ punktów na rys. 4 pozwala stwierdzić, że w przeważającej części państw europejskich dominuje ludność o wykształceniu wyższym średnim. Punkty odstające istnieją ze względu na wysokie udziały osób o wykształceniu co najwyżej niższym



\* Oznaczenia jak na rys. 1.

**Rys. 4.** Ilustracja struktury ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej z zastosowaniem wykresu trójkątnego\*

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu R

średnim, natomiast tego rodzaju dysproporcji nie ma w przypadku wykształcenia wyższego.

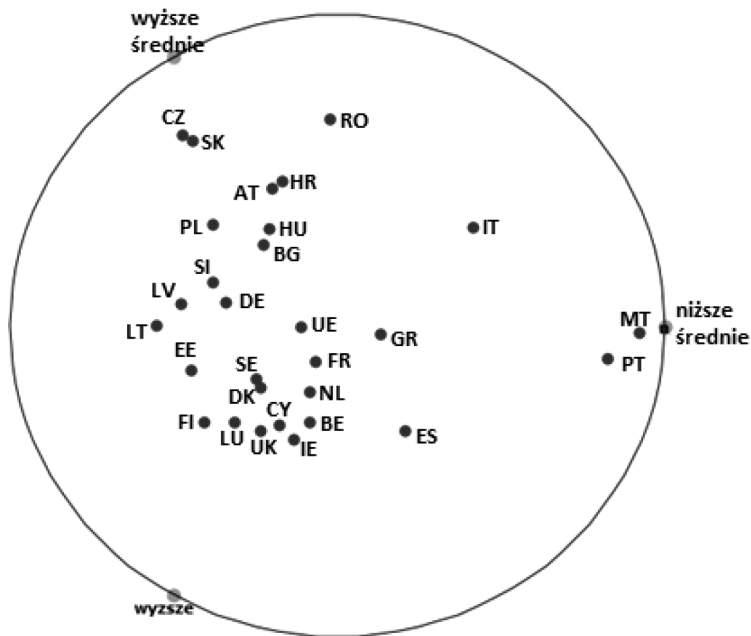
Wykres trójkątny stanowi przydatne narzędzie wizualizacji danych, ma jednak istotne wady. Po pierwsze, nie sprawdza się w przypadku struktur charakteryzujących się bardzo dużym podobieństwem. Wówczas punkty leżą bardzo blisko siebie lub nakładają się i rozpoznanie poszczególnych obiektów jest znacznie utrudnione. Po drugie, jego zastosowanie jest ograniczone do struktur trójelementowych.

W sytuacji dysponowania obiektami opisanymi za pomocą więcej niż trzech wskaźników struktury można zastosować technikę wizualizacji promieniowej (*Radial Vizualization – RadViz*) (zob. [Hoffman i in. 1997]). Jak opisują Dzemyda i in. [Dzemyda, Kurasova, Žilinskas 2012, s. 11], na obwodzie koła umieszcza się w równych odstępach  $r$  „kotwic” (reprezentujących cechy) o współrzędnych  $S_j = (s_{j1}, s_{j2})$ , a współrzędne  $Y_i = (y_{i1}, y_{i2})$  obiektów umieszczonych wewnątrz koła są wyznaczone jako:

$$Y_i = \frac{\sum_{j=1}^r S_j x_{ij}}{\sum_{j=1}^r x_{ij}}, \quad (1)$$

gdzie  $x_{ij}$  oznacza wartość  $j$ -tej cechy w obiekcie o numerze  $i$ .

Rozpatrywany zbiór danych zilustrowany za pomocą narzędzia RadViz przedstawiono na rys. 5.



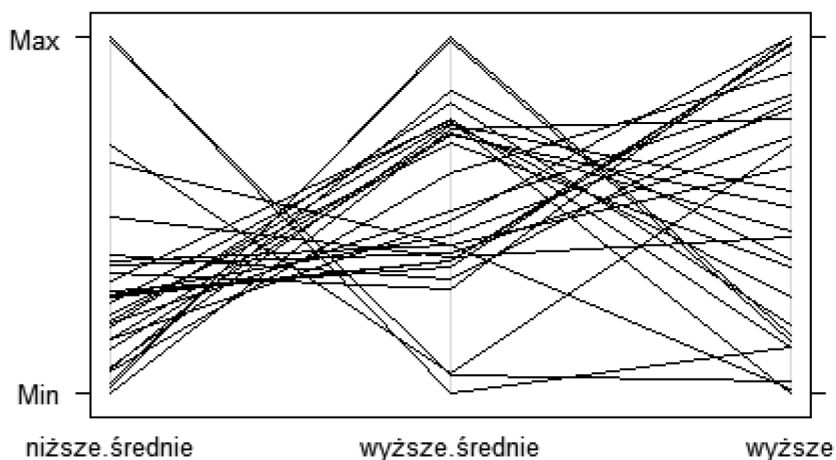
**Rys. 5.** Ilustracja struktury ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej z zastosowaniem wykresu RadViz

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu Orange [Demšar i in. 2013].

Obiekty odznaczające się dużym podobieństwem są umiejscowione blisko siebie. Obiekt o równych udziałach byłby położony dokładnie pośrodku koła, natomiast znaczna przewaga danego wskaźnika struktury skutkuje usytuowaniem punktu w bliskim sąsiedztwie odpowiadającej mu „kotwicy”. Ten sposób graficznej prezentacji danych wielowymiarowych pozwala zauważyć istniejące prawidłowości, określić skupienia obiektów podobnych i wskazać obserwacje nietypowe. Pewną wadą tego rozwiązania jest brak skali na wykresie, stąd też jego zasadnicze zastosowanie to interpretacja konfiguracji punktów względem siebie i względem „kotwic”. Z kolei niezaprzeczalną zaletą jest możliwość prezentacji struktur o wielu wymiarach.



Porównanie struktur wielowymiarowych można także przeprowadzić, wykorzystując wykres współrzędnych równoległych, na którym uwzględnione zmienne są prezentowane jako równoległe pionowe<sup>2</sup> osie wyskalowane, np. w zakresie od minimalnej do maksymalnej wartości, a poszczególne obiekty są reprezentowane przez łamaną łączącą punkty na osiach [Inselberg 1985]. Zastosowanie tego wykresu przedstawia rys. 6.



**Rys. 6.** Ilustracja struktury ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej z zastosowaniem współrzędnych równoległych

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu R.

Podobnie jak wizualizacja promieniowa, tak i wykres współrzędnych równoległych umożliwia wskazanie obserwacji nietypowych (przebieg linii łamanej odmienny od pozostałych) oraz identyfikację obiektów o podobnych strukturach (analogiczny przebieg linii łamanych). Z kolei układ punktów na osiach pozwala wnioskować o stopniu zróżnicowania poszczególnych elementów struktur.

### 3. Wizualizacja podobieństwa i braku podobieństwa struktur

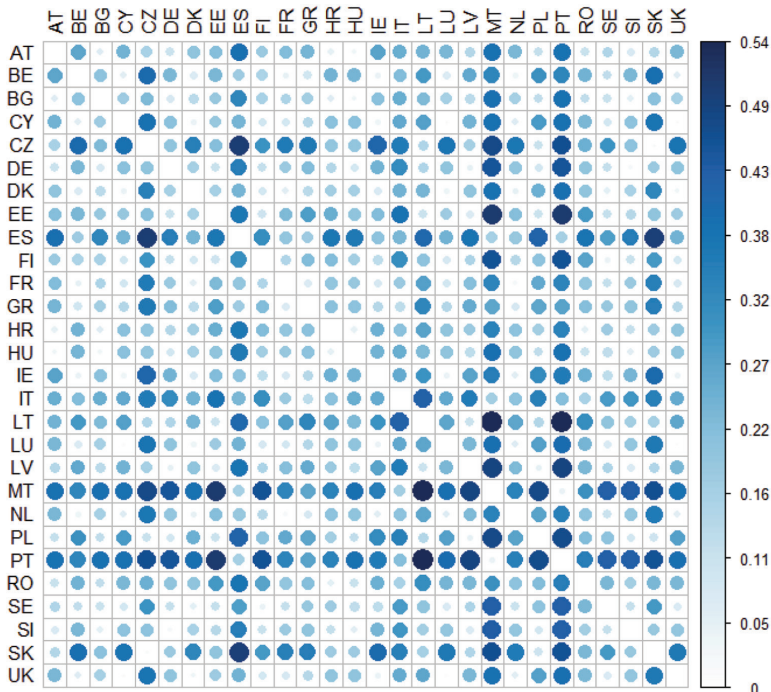
Oprócz bezpośredniej wizualizacji struktur przedmiotem zainteresowania może być także ich podobieństwo lub jego brak, które może zostać zmierzone za pomocą specyficznych miar wykorzystujących fakt, że suma składników struktury równa się jedności. Zestawienie stosowanych podejść pomiaru zawiera praca Strahl [Strahl (red.) 1998]. Do oceny braku podobieństwa struktur ludności według wykształcenia w krajach Unii Europejskiej zastosowano odległość Canberra [Strahl (red.) 1998, s. 47]:

<sup>2</sup> Układ pionowy pierwotnie proponowany przez Inselberga ma także późniejsze modyfikacje, w których osie reprezentujące zmienne przedstawiane są w poziomie.

$$d_{ij} = \frac{1}{r} \sum_{k=1}^r \frac{|p_{ik} - p_{jk}|}{p_{ik} + p_{jk}}, \quad (2)$$

gdzie:  $r$  – liczba elementów struktury,  $p_{ik}$ ,  $p_{jk}$  – wskaźniki struktury  $k$ -tej cechy w obiektach  $i$  oraz  $j$ .

Obliczając mierniki odległości dla  $n$  obiektów, otrzymuje się macierz odległości o wymiarach  $n \times n$ , której percepcja na podstawie liczb nie jest łatwa. Analizę można wspomóc, stosując techniki wizualizacji macierzy oferowane np. w pakiecie `corrplot` programu R, przeznaczonym wprawdzie przede wszystkim do odzwierciedlenia macierzy korelacji, ale umożliwiającym również graficzne ujęcie dowolnej tablicy liczb. Stopień braku podobieństwa lub podobieństwa może zostać przedstawiony za pomocą dwóch elementów łącznie: wielkości różnych symboli i natężenia koloru [Friendly 2002]. Wizualizacja macierzy odległości dla analizowanego zbioru danych z zastosowaniem kół jest przedstawiona na rys. 7.



\*Oznaczenia jak na rys. 1.

**Rys. 7.** Ilustracja macierzy braku podobieństwa struktur ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej\*

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu R.

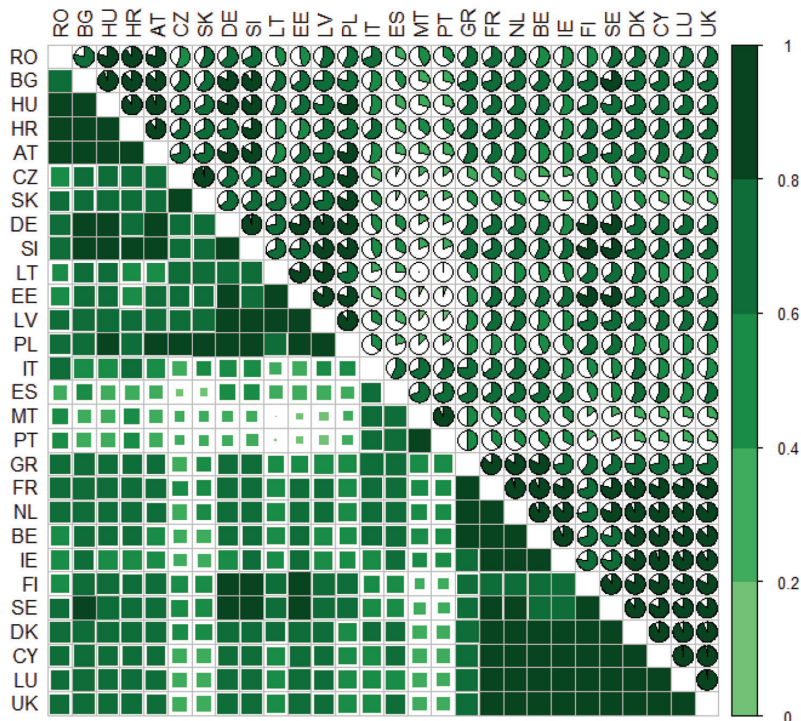
Rysunek 7 pozwala łatwo odnaleźć struktury najbardziej niepodobne do pozostałych (Portugalia, Malta, Hiszpania, Czechy) czy porównać strukturę w danym kraju z innymi, np. w porównaniu z Polską najbardziej różni się struktura na Malcie, a do odznaczających się największym podobieństwem zaliczyć można strukturę na Łotwie, Węgrzech, w Słowenii i Niemczech.

Chcąc przedstawić podobieństwo struktur, można miary odległości zamienić na współczynniki podobieństwa, np. za pomocą formuły [Brzeziński (red.) 2004, s. 509]:

$$q_{ij} = 1 - \frac{d_{ij}}{a}, \quad (3)$$

gdzie  $a$  oznacza maksymalną odległość pomiędzy dwoma różnymi obiektami.

Aby przybliżyć możliwości pakietu `corrplot` na rys. 8 ujęto graficznie macierz podobieństw z zastosowaniem kwadratów, wykresów kołowych, natężenia koloru oraz grupowania obiektów podobnych.



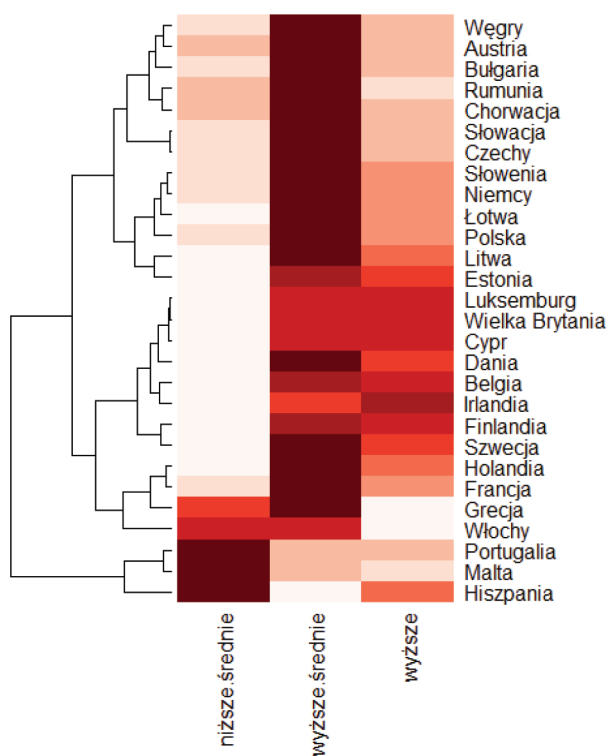
Oznaczenia jak na rys. 1\*

**Rys. 8.** Ilustracja uporządkowanej macierzy podobieństwa struktur ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej\*

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu R.

Wielkość kwadratów, stopień wypełnienia wykresów kołowych, nasycenie kolorem oraz grupowanie ułatwiają wizualną ocenę podobieństwa struktur<sup>3</sup>.

Wizualizacje zaprezentowane na rys. 7 i rys. 8 reprezentują macierz odległości i macierz podobieństwa. Innym podejściem ukazującym graficznie zgodność obiektów pod względem struktur jest połączenie rezultatów wyników grupowania hierarchicznymi procedurami aglomeracyjnymi z przedstawieniem za pomocą kolorów macierzy pierwotnych danych – ilustracja typu *heatmap* (zob. np. [Chen, Haerdle, Unwin 2008, s. 567-568]) zaprezentowana na rys. 9. Dendrogram znajdujący się po lewej stronie rysunku daje możliwość wyodrębnienia klas obiektów podobnych oraz prześledzenia kolejnych etapów procedury aglomeracyjnej. Z kolei natężenie kolorów (w wierszach) pozwala na wgląd w strukturę ludności według wykształcenia w poszczególnych krajach.



**Rys. 9.** Ilustracja grupowania struktury ludności w wieku 25-64 lat według wykształcenia w krajach Unii Europejskiej z zastosowaniem wizualizacji typu *heatmap*

Źródło: opracowanie własne na podstawie danych Eurostatu z wykorzystaniem programu R.

<sup>3</sup> Warto nadmienić, że innym sposobem graficznej prezentacji macierzy podobieństw w połączeniu z grupowaniem jest diagram Czekanowskiego, zaproponowany już na początku XX w. (zob. np. [Steczowski, Zeliaś 1981, s. 29-37]).

Dendrogram na rys. 9 sugeruje podział na pięć lub sześć skupień. Klasę najbardziej niepodobną do pozostałych, przyłączaną w ostatnim etapie aglomeracji stanowią Portugalia, Malta i Hiszpania. W grupie z Polską znajdują się Słowenia, Niemcy, Łotwa, Litwa i Estonia, czyli państwa odznaczające się bardzo wysokim odsetkiem ludności o wykształceniu wyższym średnim i stosunkowo dużym odsetkiem osób najlepiej wykształconych.

## 4. Podsumowanie

Wszystkie przedstawione narzędzia wizualizacji należą do grupy technik niewymagających redukcji wielowymiarowości, która zazwyczaj odbywa się kosztem utraty pewnej ilości informacji. Zastosowania niektórych z form nie ograniczają się jedynie do analiz struktur, choć i w tym obszarze na pewno stanowią cenne uzupełnienie prowadzonych analiz, zgodnie z powiedzeniem: „jeden obraz mówi więcej niż tysiąc słów”. Można się spodziewać, że wraz ze wzrostem liczby informacji oraz rozwojem oprogramowania, w tym darmowego, znaczenie technik graficznej prezentacji będzie coraz większe. Zaprezentowane metody mogą z powodzeniem zastąpić obszerne tabelaryczne zestawienia danych, zwłaszcza w celu poglądowego przedstawienia relacji pomiędzy strukturami wielu badanych obiektów.

## Literatura

- Aitchison J., *The Statistical Analysis of Compositional Data*, Springer, 1986.
- Brzeziński J. (red.), *Metodologia badań psychologicznych. Wybór tekstów*, Wydawnictwo Naukowe PWN, Warszawa 2004.
- Chen C.H., Haerdle W., Unwin A., *Handbook of Data Visualization*, Springer-Verlag 2008.
- Demšar J. i in., *Orange: Data mining toolbox in python*, “Journal of Machine Learning Research” 2013, 14(Aug.), s. 2349-2353.
- Dorośli w systemie edukacji formalnej: polityka i praktyka w Europie*, Komisja Europejska, wyd. polskie: Fundacja Rozwoju Systemu Edukacji, Warszawa 2011.
- Du Toit S.H.C., Steyn A.G.W., Stumpf R.H., *Graphical Exploratory Data Analysis*, Springer-Verlag, New York 1986.
- Dzemyda G., Kurasova O., Žilinskas J., *Multidimensional Data Visualization: Methods and Applications*, Springer, 2012
- Educational attainment and outcomes of education (edat)*, [http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search\\_database](http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database) [dostęp: 5.03.2014].
- Friendly M., *Corrgrams: Exploratory displays for correlation matrices*, “The American Statistician” 2002, 56.4, s. 316-324.
- Friendly M., *Visualizing Categorical Data*, SAS Institute, 2000.
- Hoffman P., Grinstein G., Marx K., Grosse I., Stanley E., *DNA visual and analytic data mining*, [w:] *Visualization '97*, Proceedings IEEE, 1997, s. 437-441.
- [http://epp.eurostat.ec.europa.eu/cache/ITY\\_SDDS/en/edat\\_esms.htm](http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/edat_esms.htm) [dostęp: 15.03.2014].
- Inselberg A., *The plane with parallel coordinates*, “The Visual Computer” 1985, 1.2, s. 69-91.
- Sarkar D., *Lattice: Multivariate Data Visualization with R*, Springer, 2008.

- Steczkowski J., Zeliaś A., *Statystyczne metody analizy cech jakościowych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1981.
- Strahl D. (red.), *Taksonomia struktur w badaniach regionalnych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 1998.
- Suchecki B. (red.), *Ekonometria przestrzenna: metody i modele analizy danych przestrzennych*, Wydawnictwo C.H. Beck, Warszawa 2010.

## **APPLICATION OF SELECTED FORMS AND TECHNIQUES OF DATA VISUALIZATION IN ANALYZING SOCIO-ECONOMIC STRUCTURES**

**Summary:** The paper presents a selection of visualization methods which may be useful for the analyses and comparisons of socio-economic structures. The work is focused on the techniques that do not require dimensionality reduction, which is usually accompanied by the loss of some information from the dataset. The methods of graphical presentation of structures and illustration of their similarities and dissimilarities are considered as well as strengths and weaknesses of various solutions are outlined. The discussion is based on data concerning the structures of education attainment in the European Union countries.

**Keywords:** analysis of socio-economic structures, data visualization, education attainment structure.