

Ograniczenia analizy skupień metodą k-średnich w tworzeniu typologii obiektów

Jerzy Wierziński

Łatwość, z jaką za pomocą odpowiedniego programu komputerowego można dokonać skomplikowanych analiz danych, nie idzie niestety w parze ze zrozumieniem przez badaczy istoty wykonywanych analiz. Celem artykułu jest pokazanie ograniczeń analizy skupień (cluster analysis) metodą k-średnich, używanej powszechnie w badaniach marketingowych do segmentacji rynku, w psychologii do wykrywania różnic indywidualnych, stosowanej także do kategoryzacji krajów. Tok wywodu został zilustrowany wynikami własnych analiz wskaźników religijności w badaniach ISSP prób reprezentatywnych pochodzących z 32 krajów. Pokazano, że na podstawie tych samych danych możemy stworzyć za pomocą analizy skupień różne, spełniające kryteria statystyczne, kategoryzacje. Aby stworzyć użyteczną merytorycznie klasyfikację, trzeba testować różne warianty rozwiązań, pamiętając, że ważniejsza od istotności statystycznej, którą w przypadku analizy skupień można bardzo łatwo uzyskać, jest istotność merytoryczna.

1. Analiza skupień (cluster analysis) metodą k-średnich

Analiza skupień (cluster analysis) metodą k-średnich jest powszechnie używana w badaniach marketingowych do segmentacji rynku (Kacprzak-Chońska 2009), w psychologii do wykrywania różnic indywidualnych (Lewicka 2004), stosowana jest także do kategoryzacji krajów (Johnston 2005). Celem tego artykułu jest zwrócenie uwagi na jej ograniczenia. Opis analizy skupień pokażemy na konkretnym przykładzie kategoryzacji dotyczącej stopnia religijności, ale wszystkie opisywane własności tej metody mają charakter ogólny.

Zmienne opisujące obiekty to odpowiedzi na 4 pytania z sondażu ISSP¹ przeprowadzonego w latach 1998/1999 na próbach reprezentatywnych w 32 krajach (zob. tabela 1 na s. 235):

- rel1: „Jak często modlisz się?” z odpowiedziami na jedenastopunktowej skali odpowiedzi od <nigdy> do <kilka razy dziennie>,
- rel2 „Jak określiłbyś swoją religijność?” z odpowiedziami na siedmiopunktowej skali odpowiedzi od <bardzo religijna> do <wyjątkowo nieregelijna>,

- rel3 „Jak często bierzesz udział w praktykach religijnych?” z odpowiedziami na dziewięciopunktowej skali odpowiedzi od <nigdy> do <kilka razy w tygodniu>,
- rel4 „Jak duże masz zaufanie w stosunku do Kościołów i organizacji wyznaniowych?” z odpowiedziami na sześciopunktowej skali odpowiedzi od <całkowite zaufanie> do <w ogóle nie mam zaufania>.

Odpowiedzi na te pytania mogą być uznane za wskaźniki zmiennej teoretycznej <religijność> (por. Wieczorkowska i Wierziński 2007). W dalszych analizach odpowiedzi na pytanie 2. i 4. zostały zrekodowane tak, aby wyższe wartości zmiennej świadczyły o wyższym natężeniu zmiennej teoretycznej.

Wskaźniki zmiennej teoretycznej (ukrytej, latentnej) można podzielić na:

- **wskaźniki proste**, którymi są np. zakodowane numerycznie odpowiedzi na pytania,
- **indeksy złożone**, które budujemy z kilku, kilkunastu lub czasem nawet kilkudziesięciu wskaźników prostych – do tego celu najczęściej używane są średnie arytmetyczne (po analizie jednorodności skali za pomocą α Cronbacha) lub wyniki czynnikowe po wykonaniu analizy czynnikowej.

W naszym przykładzie moglibyśmy zbudować charakteryzujący się wysoką jednorodnością (α Cronbacha = 0,71 dla 35 186 respondentów) indeks złożony z 4 pytań, ponieważ tworzą one jeden czynnik wyjaśniający ponad 61% wariancji zmiennych. Uśrednianie 4 wskaźników prostych zaciera jednak ich zróżnicowanie.

Jeżeli nie chcemy tego robić, musimy uciec się do bardziej skomplikowanej analizy, której przykład będzie omówiony w dalszej części tekstu.

Drugim kryterium podziału wskaźników jest jednostka analizy. Wskaźnik może opisywać obiekt pojedynczy (np. respondenta) lub złożony (np. grupę respondentów). Możemy je nazwać odpowiednio wskaźnikami **jednostkowymi** v. **agregatowymi**. Brak określenia typu wskaźnika oznacza domyślnie, że używano prostych wskaźników jednostkowych.

W naszym przykładzie używać będziemy zarówno prostych wskaźników agregatowych (średnie odpowiedzi na 4 pytania w 32 krajach), jak i prostych wskaźników jednostkowych (odpowiedzi na 4 pytania ponad 32 tysięcy respondentów).

Aby zwiększyć czytelność tego metodologicznego tekstu, skoncentrujemy się na początku na kategoryzacji krajów na podstawie wskaźników prostych. W naszym przypadku dane wejściowe mają postać macierzy 32 (kraje) \times 4 (zmienne).

Celem analizy skupień metodą k-średnich jest:

- jednoczesne porównywanie obiektów opisanych na wielu wymiarach,
- zredukowanie dużej liczby obiektów do kilku podstawowych typów, które mogą być traktowane jako obiekty dalszej analizy.

Terminy: skupienia, kategorie, typy będą używane w tym tekście zamiennie.

Analiza skupień metodą k-średnich (Aldenderfer i Blashfield 1984; Bailey 1994) grupuje obiekty – w naszym przypadku kraje – w kategorie, które powinny być zróżnicowane: minimalnie wewnątrz skupienia i maksymalnie między skupieniami. Każdy kraj przedstawiony jest jako punkt o czterech współrzędnych (średnie odpowiedzi na 4 pytania w 32 krajach). Możemy więc te dane przedstawić w czterowymiarowej przestrzeni.

Zadaniem analizy jest odkrycie w tej przestrzeni skupień punktów (czyli obiektów leżących blisko siebie). Podobieństwo krajów przekłada się na bliskość odpowiadających im punktów w przestrzeni, co możemy ocenić, licząc odległości między nimi. Są to zwykle odległości euklidesowe, będące pierwiastkiem kwadratowym z sumy kwadratów różnic współrzędnych na poszczególnych wymiarach. **Punkt centralny skupienia** nazywany jest jego **centroidem (środkiem ciężkości)**. Celem analizy jest odkrycie skupień punktów i ich opisanie poprzez wskazanie środka ciężkości (centroidu) obiektów należących do skupienia oraz ich odległości od środka ciężkości. Im mniejsza jest odległość punktu od centroidu, tym lepiej dany kraj pasuje do skupienia.

W naszym przykładzie danymi wejściowymi są opisy 32 krajów na czterech wymiarach. Dane wyjściowe to **współrzędne centroidów**, informacja o przypisaniu analizowanych obiektów do skupień oraz **odległości danego obiektu od centroidu skupienia**, do którego został zakwalifikowany. Dodatkowo liczone są odległości między centroidami.

Podstawowym problemem jest wybór liczby skupień, ponieważ badacz musi sam określić liczbę kategorii, którą chce otrzymać. Nie istnieje jedna metoda wyboru liczby kategorii w przypadku stosowania analizy k-średnich. Naszym celem jest znalezienie:

- najmniejszej liczby skupień, przy której konfiguracja skupień jest najlepszym z możliwych dopasowań do danych (**kryterium istotności statystycznej**),
- takiej konfiguracji, która jest najłatwiej interpretowalna i ma samodzielne znaczenie (**kryterium istotności merytorycznej**).

Im większa liczba skupień, tym łatwiej spełnić założenie o wystarczającym dopasowaniu otrzymanej konfiguracji wyjściowej do danych wejściowych. Jak pokazałem to w innej pracy (Wierziński 2009), podjęcie decyzji o liczbie skupień może ułatwić analiza rozkładu odległości obiektów od centroidów przy różnej liczbie kategorii. Na podstawie analizy rozkładu odległości od centroidów można wykonać **wykres osypiska** analogiczny do tego, który wykonujemy w analizie czynnikowej (Wieczorkowska i Wierziński 2007), aby podjąć decyzję o liczbie czynników. Trzeba jednak pamiętać, że najważniejszym kryterium wyboru optymalnej liczby skupień (kategorii) jest sensowność interpretacji różnych rozwiązań. Znaczenie merytoryczne różnych kategorii jest równie ważnym wskaźnikiem, który należy brać pod uwagę przy ustalaniu optymalnej liczby wymiarów, jak i różne statystyczne testy rozwiązania. Zgadzam się całkowicie z Torgersonem

(1958), który mówił, że w skalowaniu wielowymiarowym, które podobnie jak analiza skupień opiera się na programach iteracyjnych, interpretowalność i stabilność rozwiązania powinny być podstawowym kryterium w obieraniu optymalnej wymiarowości w skalowaniu – w przypadku analizy skupień odnosi się to do optymalnej liczby kategorii.

Aby ułatwić ilustrację kolejnego problemu, skoncentrujemy się w dalszych analizach na rozwiązaniach wymuszających 3 skupienia. Przeprowadzona analiza eksploracyjna podzieliła 32 kraje na 3 skupienia o liczebności odpowiednio 5, 17 i 10 obiektów.

Czy możemy uznać to spełniające kryteria statystyczne rozwiązanie za satysfakcjonujące? Zanim przystąpimy do merytorycznej interpretacji rozwiązania, musimy się zastanowić, czy otrzymany wynik jest wyznaczony jednoznacznie, czy też jest tak, że przy powtarzaniu analiz otrzymamy zupełnie inne rozwiązania.

Aby zrozumieć istotę otrzymanego rozwiązania, musimy przyjrzeć się bliżej zastosowanej procedurze. Program szuka rozwiązania **za pomocą kolejnych iteracji**. Wyznacza początkowe (startowe) centroidy i sprawdza, jak są one dopasowane do obiektów, następnie zmienia centroidy tak, aby były lepiej dopasowane, i robi to tak długo, aż żadna następna zmiana nie poprawia wskaźnika dopasowania (SES), którym jest suma odległości wszystkich obiektów od centroidu skupienia, do którego obiekt został zaklasyfikowany.

Jednym z niebezpieczeństw jest problem lokalnego minimum (por. np. Wiczorkowska 1987). Program używa metod minimalizowania funkcji SES, czego rezultatem może być osiągnięcie jedynie tzw. **lokalnego minimum**, a nie oczekiwanego prawdziwego minimum. Zachowanie programu można obrazowo porównać do spacerowania po okolicy w celu znalezienia najniższego punktu. Za punkt minimalny uznajemy taki, który zawsze będzie położony najniżej – bez względu na to, w którym kierunku się udamy. W programach iteracyjnych oznacza to, że każdy następny krok (kolejna iteracja) zwiększa wskaźnik dopasowania – wartość funkcji SES. Może się jednak zdarzyć, że ten pozornie wyglądający jak najniższy punkt jest płytką doliną, tymczasem obok znajduje się depresja, której nie mamy szansy odkryć, ponieważ usatysfakcjonowani skończyliśmy poszukiwania w dolinie. Aby upewnić się, że nie mamy do czynienia z takim właśnie przypadkiem, należy uruchomić program kilkakrotnie z różnymi konfiguracjami wejściowymi. Jeżeli rozwiązania są identyczne, możemy sądzić, że jest to prawdziwe minimum.

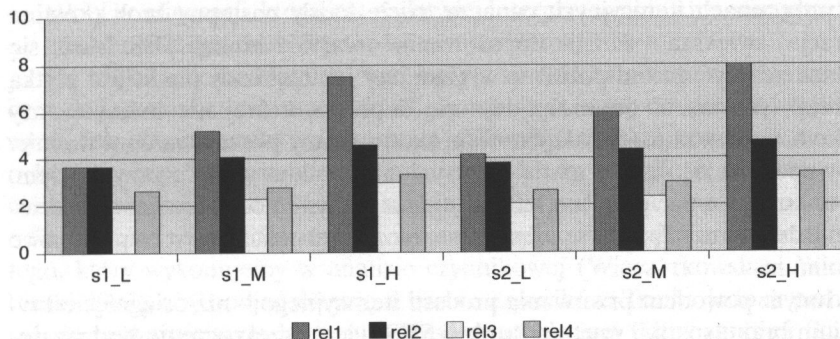
Innym powodem przerwania procesu iteracyjnego – niż osiągnięcie kryterium braku spadku wartości funkcji SES – jest przekroczenie liczby założonych iteracji.

Aby sprawdzić stabilność otrzymanego rozwiązania, zmieniono posortowanie zbioru wejściowego i okazało się, że ta niemająca przecież merytorycznie żadnego znaczenia zmiana uporządkowania zbioru (sortowanie 1. v. sor-

towanie 2.) spowodowała zmianę przypisania krajów do poszczególnych skupień. Dlaczego? Otóż kluczowym punktem we wszystkich programach iteracyjnych jest wybór centroidów startowych. Algorytm używany przez popularny pakiet statystyczny SPSS (Hartigan 1975) szukanie centroidów startowych zaczyna od wybrania **pierwszych** trzech obiektów w zbiorze (stąd sortowanie zbioru ma wpływ na wyniki) – są to centroidy początkowe C1, ..., C3. Następnie dla czwartego kraju, nazwijmy go X, liczy odległość od C1, ..., C3. Jeśli odległość między X a najbliższym mu centroidem jest większa niż dystans między dwoma najbliższymi centroidami, wtedy X zastępuje C1 lub C2, w zależności od tego, który jest bliższy. Jeżeli X nie spełnia tego kryterium, program sprawdza, czy jest on oddalony od C2 bardziej niż C2 od jakiegokolwiek innego centroidu. Jeżeli tak jest, to najbliższy centroid jest zastąpiony przez X. Ta procedura jest powtarzana **raz** dla wszystkich obiektów w zbiorze. Po jej zakończeniu startowe centroidy są już ustalone.

Przy sortowaniu 1. jako centroidy startowe zostały wybrane Filipiny, Niemcy Wschodnie i Chile. Przy sortowaniu 2. zbioru wejściowego na liście wybranych przez program centroidów startowych miejsce Japonii zajął Cypr. Przy zmienionych centroidach startowych analiza zakończyła się podziałem na skupienia zawierające odpowiednio 8, 14 i 10 obiektów. Dane wejściowe były takie same, ale inne posortowanie zbioru spowodowało otrzymanie innej klasyfikacji. Wartości centroidów finalnych przedstawione są na rysunku 1. Przy obu rozwiązaniach profile wskaźników są podobne, ale aż 10 z 32 krajów zmieniło skupienie w obu analizach. Nie musi być to częste, ale może się zdarzyć.

Warto podkreślić, że wpływ sortowania (mogącego wpłynąć na wybór konfiguracji startowej) maleje wraz z liczbą kategoryzowanych obiektów i rośnie wraz z liczbą założonych skupień.



Uwaga: Skupienia o wysokich, średnich i niskich wartościach wskaźników oznaczono odpowiednio symbolami: H, M, L.

Rys. 1. Wartości centroidów finalnych dla trzech skupień przy dwóch (s1 v. s2) różnych sortowaniach zbioru wejściowego. Źródło: opracowanie własne.

Kolejnym problemem są różnice w otrzymanym rozwiązaniu w zależności od tego, czy zmienne wejściowe są wystandaryzowane czy nie. Gdy wariancje zmiennych nie różnią się znacznie, standaryzacja nie powinna wpłynąć na konfigurację finalną. W przeciwnym wypadku otrzymamy różne rozwiązania i pojawi się pytanie, które z rozwiązań trafniej odzwierciedla „rzeczywistość”.

Podsumowując, warto pamiętać, że analiza skupień metodą k-średnich może być źródłem hipotez, ale nie dowodem na istnienie kategorii, o ile badacz nie udowodni, że otrzymane przez niego rozwiązanie jest odporne na zmianę konfiguracji startowej i wybór liczby skupień. To ostatnie oznacza, że „przemieszczanie” się krajów między skupieniami w rozwiązaniach różniących się liczbą kategorii daje się sensownie interpretować.

2. Testowanie hipotez za pomocą analizy skupień

Zdarza się, że początkujący badacze uważają, że prawdziwa praca naukowa polega wyłącznie na testowaniu hipotez. Efektem tego jest dopisywanie sztucznych hipotez do dowolnej aktywności badawczej. Ci, którzy tworzą klasyfikację, formułują hipotezę: „Istnieje możliwość uporządkowania zbioru obiektów”. Mendelejew powinien więc być napisać, że jego praca nad układem okresowym pierwiastków chemicznych polegała na testowaniu hipotezy, że da się takie uporządkowanie przeprowadzić. Jest to absurdalny pomysł, ponieważ, gdyby mu się nie udało – tak jak jego poprzednikom – nie oznaczałoby to, że nie można tego zrobić. Podobnie u badaczy stosujących analizę skupień metodą k-średnich można znaleźć hipotezę sformułowaną np. w następujący sposób: „Rynek młodego konsumenta jest heterogeniczny pod względem wybranych czterech zmiennych społeczno-kulturowych i istnieje na nim homogeniczne wewnętrznie segmenty”.

Przypomnijmy, że hipotezą empiryczną nazywamy zdanie, które może zostać sfalsyfikowane w wyniku analizy danych. W przypadku przytoczonej wyżej hipotezy trudno sobie wyobrazić, jakie wyniki empiryczne mogłyby ją sfalsyfikować. Za pomocą analizy skupień dostaniemy żadaną liczbę kategorii nawet wtedy, gdy zmienne klasyfikujące mają identyczne średnie i wariancje, czy też charakteryzują się pełnym skorelowaniem, a więc brak zróżnicowania rynku moglibyśmy stwierdzić tylko wtedy, gdyby zmienne wejściowe miały zerową wariancję, a więc były stałymi, a nie zmiennymi.

Zdecydowanie zachęcam badaczy, którzy chcą analizować swoje dane o „nieznanej strukturze”, aby najpierw zobaczyli rozwiązania oferowane przez pakiet statystyczny dla danych spreparowanych (*made-up data*), takich jak te opisane powyżej (silne skorelowanie zmiennych, bardzo mała wariancja, brak różnic w średnich i odchyleniach standardowych).

Ze względu na opisane wyżej problemy analiza skupień mogłaby służyć testowaniu hipotez, gdyby centroidy startowe zostały wybrane przez badacza i zostałyby wybrana opcja w programie SPSS <CLASSIFY>. Choć finalne

centroidy mogą się w tym przypadku także różnić od założonych, ale zaletą tej opcji jest fakt, że otrzymane rozwiązanie nie zależy od sortowania zbioru wejściowego.

Załóżmy, że pod wpływem wyników analizy eksploracyjnej i rozważań teoretycznych (Giddens 2005; Marody i Mandel 2007) chcemy wyróżnić kraje o poziomie religijności:

- 1) wysokim (wysokie wartości wszystkich zmiennych),
- 2) niskim (niskie wartości wszystkich zmiennych),
- 3) umiarkowanym, wskazującym na postępującą sekularyzację przejawiającą się silniej w spadku religijności obrzędowej (niskie wartości wskaźnika partycypacji w uroczystościach religijnych i niskie zaufanie do instytucji religijnych) niż religijności duchowej (średnie wartości wskaźników rel1 i rel2).

Problemem jest wybór wartości centroidów startowych. Można wybrać konkretne kraje, ale można też wpisać wartości centroidów teoretycznych, nie reprezentowanych dokładnie przez żaden kraj (Wierziński 2009). Aby zoperacjonalizować wysokie, niskie i średnie wartości zmiennych wejściowych, możemy np. wpisać liczby odpowiadające wynikom równym odpowiednio $z = 1$, $z = -1$, $z = 0$ (gdzie z oznacza wystandaryzowane wartości zmiennych wejściowych). Trzeba jednak **dokładnie oglądać rozkłady zmiennych**, ponieważ proponowane rozwiązanie jest akceptowalne w naszym przykładzie dla wskaźników agregatowych, których rozkłady były w miarę symetryczne, ale nie dla wskaźników prostych, ponieważ np. rozkład wskaźnika partycypacji w próbie ponad 32 tysięcy respondentów jest wyjątkowo prawoskośny (aż 57% z nich stwierdziło, że nigdy nie chodzi do kościoła).

W tabeli 1 w kolumnach A, B, C, D przedstawione są średnie wartości wskaźników dla poszczególnych krajów, w kolumnie E zaś symbol skupienia, do którego dany kraj został zakwalifikowany. Kategorię o wysokim, niskim i średnim poziomie religijności oznaczono odpowiednio symbolami H, L, M.

Kategoryzowanie krajów sprzyja tworzeniu stereotypów (nadmiernych uogólnień) narodowych. Warto więc sprawdzić, jakie jest zróżnicowanie profili religijności w poszczególnych krajach. W tym celu przeprowadzono confirmacyjną analizę skupień ponad 32 tysięcy respondentów z centroidami startowymi o wybranych wartościach na tych samych zasadach, jak to było w przypadku krajów (konkretne wartości były jednak różne, ponieważ rozkład wskaźników agregatowych różni się od rozkładu wskaźników jednostkowych).

Trzy wyodrębnione skupienia w wyniku tej analizy podzieliły respondentów prawie równo. Do skupienia o wysokiej religijności zakwalifikowano około 38%, do pozostałych zaś po około 31%.

W kolumnach F, G, H tabeli 1 pokazano, jaki procent respondentów z danego kraju należy do kategorii o wysokim (H), niskim (L) bądź średnim (M) poziomie religijności. Tylko w przypadku Filipin prawie 91% respon-

Nr	Kraj	A	B	C	D	E	F	G	H
		rel1	rel2	rel3	rel4	skupienie	H	L	M
1	Norwegia	4,0	4,1	2,2	2,9	L	21,8	40,1	38,1
2	Szwecja	3,3	3,5	2,2	2,6	L	16,3	51,5	32,1
3	Dania	3,3	3,8	2,2	3,0	L	18,8	55,6	25,6
4	Hiszpania	5,4	4,2	1,8	2,9	M	36,3	30,5	33,2
5	Portugalia	6,8	4,8	2,1	3,3	H	54,7	14,5	30,7
6	Cypr	5,9	5,1	1,8	3,2	M	37,6	12,0	50,4
7	Słowenia	4,4	3,9	1,9	2,6	M	26,7	37,8	35,5
8	Węgry	5,5	4,1	1,6	3,1	M	38,4	32,6	28,9
9	Czechy	3,6	3,6	1,9	2,4	L	21,0	52,2	26,8
10	Słowacja	5,7	4,4	3,0	3,1	M	48,2	25,1	26,7
11	Polska	7,9	5,0	2,6	3,3	H	66,8	5,8	27,4
12	Łotwa	4,5	4,0	1,6	3,1	M	25,8	33,8	40,4
13	Bułgaria	3,9	4,1	2,6	2,6	L	23,8	38,7	37,5
14	Rosja	3,4	3,9	1,1	3,0	L	19,7	53,9	26,4
15	Niemcy Wsch.	2,4	2,3	1,7	2,2	L	12,7	72,1	15,2
16	Niemcy Zach.	4,8	3,8	2,8	2,7	M	32,3	34,4	33,3
17	Austria	5,7	4,4	2,7	2,7	M	40,6	22,1	37,3
18	Włochy	6,7	4,4	2,9	3,0	H	53,9	16,6	29,6
19	Holandia	4,8	4,2	2,0	2,7	M	35,6	39,5	24,9
20	Francja	4,1	3,5	2,0	2,4	L	23,3	47,0	29,6
21	Szwajcaria	6,0	3,7	2,6	2,9	M	39,2	22,6	38,2
22	Wielka Brytania	4,6	4,0	2,2	2,7	M	28,8	33,2	38,0
23	Irlandia	7,4	4,7	2,6	3,0	H	60,9	6,4	32,7
24	Irlandia Północna	6,8	4,3	2,9	3,1	H	55,0	18,2	26,8
25	USA	7,8	4,8	3,4	3,3	H	66,8	11,1	22,0
26	Kanada	5,4	4,1	2,5	2,8	M	35,8	28,8	35,4
27	Australia	4,6	4,0	2,4	2,9	M	33,8	38,9	27,2
28	Nowa Zelandia	4,9	4,1	2,7	2,8	M	33,0	33,6	33,4
29	Filipiny	9,4	5,1	4,0	3,9	H	90,6	0,3	9,1
30	Japonia	5,3	3,5	1,9	2,1	M	25,6	24,3	50,0
31	Chile	7,3	4,1	3,2	3,6	H	61,2	14,9	23,9
32	Izrael	4,7	3,7	2,5	2,6	M	30,8	43,3	25,9

Uwaga: objaśnienia kolumn A-H znajdują się w tekście.

Tab. 1. Lista krajów biorących udział w badaniu ISSP. Źródło: opracowanie własne.

dentów zostało zakwalifikowanych do tej samej kategorii. W wielu krajach podział respondentów między trzy kategorie jest prawie równy, co pozwala sądzić, że powinno być trudno zakwalifikować ten kraj na podstawie wskaźnika agregatowego.

Zaprezentowane analizy są jedynie fragmentem tego, co należałoby zrobić, gdyby naszym zadaniem było stworzenie klasyfikacji krajów pod względem religijności. Potrzebna jest np. analiza rozwiązań o wyższej liczbie skupień, rozważenie argumentów za standaryzacją zmiennych wejściowych i przeciw niej, sprawdzenie wpływu kraju odstającego od reszty (Filipiny) na kategoryzację pozostałych.

Szybkość, z jaką można za pomocą odpowiedniego programu komputerowego dokonać klasyfikacji (kategoryzacja ponad 32 tysięcy obiektów trwa mniej niż minutę), nie przekłada się na łatwość uzyskiwania ważnych merytorycznie wyników. Trzeba pamiętać, że na podstawie tych samych danych możemy stworzyć za pomocą analizy skupień bardzo różne, spełniające kryteria statystyczne kategoryzacje. Aby stworzyć użyteczną merytorycznie klasyfikację, trzeba spędzić wiele dni, jeśli nie tygodni, testując różne warianty rozwiązań i pamiętając, że ważniejsza od istotności statystycznej jest istotność merytoryczna.

Informacje o autorze

Dr hab. Jerzy Wierziński, prof. UW – Kierownik Zakładu Metod Matematycznych i Statystycznych Zarządzania, Prodziekan Wydziału Zarządzania Uniwersytetu Warszawskiego.

E-mail: wierzbinski@mail.wz.uw.edu.pl.

Przypisy

- ¹ International Social Survey Program (ISSP) to stały międzynarodowy program corocznych badań porównawczych. Celem programu jest przeprowadzanie raz do roku przez wszystkie kraje członkowskie badania na określony temat, według tej samej metodologii, z zastosowaniem identycznego zestawu pytań. Dystrybucją danych ISSP dla ośrodków naukowych w Polsce zajmuje się, wraz z Archiwum Danych Społecznych, Instytut Studiów Społecznych UW. W artykule wykorzystano wyniki badania z 1998/1999 r. (ponad 32 tysięcy respondentów z 32 krajów).

Bibliografia

- Aldenderfer, M.S. i R.K. Blashfield. 1984. *Cluster analysis*, Thousand Oaks–London–New Delphi: Sage Publications.
- Bailey, K.D. 1994. *Typologies and Taxonomies. An Introduction to Classification Techniques*, Thousand Oaks–London–New Delphi: Sage Publications.
- Giddens, A. 2005. *Socjologia*, Warszawa: Wydawnictwo Naukowe PWN.
- Hartigan, J.A. 1975. *Clustering Algorithms*, New York: John Wiley & Sons.
- Johnston, M. 2005. *Syndromes of Corruption: Wealth, Power, and Democracy*, Cambridge: Cambridge University Press.
- Kacprzak-Chońska, A. 2009. *Spoleczno-kulturowe determinanty segmentacji rynku młodzieżowego*, Warszawa: Wydział Zarządzania, Uniwersytet Warszawski, niepublikowana praca doktorska.

- Lewicka, M. 2004. Gdybym miał milion... Marzenia Polaków u progu XXI wieku. *Kolektyw Psychologiczne*, nr 12, s. 218–237.
- Marody, M. i S. Mandes. 2007. Religijność a tożsamość narodowa Polaków, w: Marody, M. (red.) *Wymiary życia społecznego. Polska na przełomie XX i XXI wieku. Wydanie nowe*, Warszawa: Wydawnictwo Naukowe Scholar.
- Torgerson, W.S. 1958. *Theory and Methods of Scaling*, New York: John Wiley & Sons.
- Wieczorkowska, G. 1987. *Skalowanie wielowymiarowe jako metoda badania percepcji*, Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Wieczorkowska, G. i J. Wierziński. 2007. *Statystyka. Analiza badań społecznych*, Warszawa: Wydawnictwo Naukowe Scholar.
- Wierziński, J. 2009. *Badanie zaufania do organizacji: problemy metodologiczne*, Warszawa: Wydawnictwa Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego.