

Justyna Brzezińska

Akademia Ekonomiczna w Katowicach

ANALIZA I WIZUALIZACJA RYNKU STACJI RADIOWYCH NA ŚLĄSKU

1. Wstęp

Metody klasyfikacji znajdują bardzo szerokie zastosowanie w praktyce, pozwalając badaczowi na grupowanie i klasyfikowanie obiektów, dostarczając istotnych informacji o badanym zjawisku lub zachowaniach na rynku.

W niniejszej pracy przeprowadzono analizę rynku radiowego na Śląsku. W części pierwszej scharakteryzowano dane dotyczące rynku radiowego. Klasyfikacji obiektów dokonano za pomocą metody *k*-średnich, a następnie jej wyniki zweryfikowano z wynikami klasyfikacji otrzymanej przy użyciu skalowania wielowymiarowego. Przeprowadzenie badania ma na celu wyodrębnienie jednorodnych grup stacji radiowych na Śląsku, a co za tym idzie – poznanie rynku radiowego i jego ocenę przez respondentów biorących udział w badaniu. Wyniki wykorzystanych metod przedstawiono w postaci graficznej na wykresie rozrzutu oraz w postaci mapy percepcji. Część ostatnia zawiera wnioski oraz interpretację otrzymanej klasyfikacji, tj. charakterystykę powstałych grup i należących do nich stacji radiowych.

2. Charakterystyka danych dotyczących rynku stacji radiowych

Analizie poddanych zostanie czternaście obiektów, przedstawiających stacje radiowe opisywane za pomocą dziesięciu zmiennych. Obiektami w przeprowadzonym badaniu są stacje radiowe: RMF (93,0 MHz), Blue (94,5 MHz), ZET (95,7 MHz), Plus (96,2 MHz), Eska (99,1 MHz), „Trójka” (99,7 MHz), Katowice (102,2 MHz), Bis (105,6 MHz), Em (107,6 MHz), Tok (97,4 MHz), Planeta (95,1 MHz), Mega (92,3 MHz), Karolina (91,2 MHz), „Jedynka” (97,9 MHz).

Zmiennymi są cechy charakteryzujące każdą ze stacji radiowych, tj.: liczba reklam, dostęp słuchaczy na antenę, wiek słuchaczy, rodzaj emitowanej muzyki, częstotliwość informacji, liczba informacji politycznych, poziom prowadzonych audycji, promocja i reklama stacji, pora nadawania audycji i liczba serwisów edukacyjnych i kulturalnych.

W badaniu wzięło udział 200 respondentów. Skalą dyferencjału semantycznego w niniejszym badaniu jest 7-stopniowa skala Likerta pozwalająca respondentowi na wybór kategorii od „1” do „7” względem wszystkich badanych zmiennych dla każdego obiektu (tab. 1).

Tabela 1. Skala dyferencjału semantycznego wykorzystana w badaniu ankietowym

Nazwa stacji radiowej		1	2	3	4	5	6	7	
Mała liczba reklam									Duża liczba reklam
Łatwy dostęp słuchaczy na antenę									Trudny dostęp słuchaczy na antenę
Adresowane do osób młodych									Adresowane do osób starszych
Muzyka dla koneserów									Muzyka popularna
Częste informacje									Rzadkie informacje
Dużo informacji politycznych									Mało informacji politycznych
Wysoki poziom prowadzenia audycji									Niski poziom prowadzenia audycji
Częsta reklama i promocja stacji									Brak reklamy i promocji stacji
Odpowiednia pora nadawania audycji									Nieodpowiednia pora nadawania audycji
Dużo serwisów edukacyjnych i kulturalnych									Mało serwisów edukacyjnych i kulturalnych

Źródło: opracowanie własne.

Tabela 2. Macierz średnich odpowiedzi uzyskanych w badaniu własnym ($n = 14$ obiektów, $m = 10$ zmiennych)

	Reklama	Antena	Wiek	Muzyka	Informacje	Polityka	Poziom	Promocja	Pora	Kultura
RMF	5,11	3,83	3,21	5,42	2,99	3,42	3,34	2,66	3,20	4,75
Blue	4,13	3,64	3,51	4,71	3,76	4,37	3,97	3,85	3,52	4,81
ZET	5,23	4,07	3,35	5,35	3,10	3,28	3,65	2,64	3,27	4,84
Plus	3,60	3,75	4,09	4,32	4,05	4,28	3,82	4,39	3,56	3,86
Eska	4,35	2,96	2,15	5,81	4,62	5,21	3,98	3,15	3,07	5,33
„Trójka”	3,65	3,84	4,41	3,01	3,13	3,05	2,47	4,14	3,22	2,81
Katowice	4,22	3,26	4,63	4,12	3,28	3,58	3,74	4,22	3,54	3,48
Bis	3,60	3,53	3,07	3,60	3,93	4,93	3,73	5,07	2,60	2,80
Em	3,74	3,65	4,52	4,03	3,74	4,35	3,81	3,90	3,16	3,35
Tok	3,20	3,60	4,55	4,45	2,65	2,85	2,80	3,95	3,65	3,65
Planeta	3,33	2,67	1,78	5,11	4,83	5,56	3,14	3,28	2,36	5,39
Mega	4,38	4,38	3,75	5,38	4,50	5,50	4,75	4,63	4,13	4,63
Karolina	3,60	3,40	5,60	3,87	4,07	5,60	4,07	4,13	2,27	5,33
„Jedynka”	1,83	3,83	5,17	2,00	3,83	5,17	2,50	6,67	2,17	3,83

Źródło: opracowanie własne.

Aby trójwymiarową kostkę danych (obiekty, zmienne, wartość skali Likerta) zredukować do zbioru dwuwymiarowego, należy obliczyć średnie wartości odpowiedzi 200 respondentów, którzy wzięli udział w badaniu (tab. 2).

Następnie dla tego zbioru danych obliczono macierz odległości euklidesowych pomiędzy obiektami (tab. 3).

Tabela 3. Macierz odległości między obiektami

Macierz odległości														
Użyto - Odległość euklidesowa														
	1:RMF	2:Blue	3:ZET	4:Plus	5:Eska	6:Trójka	7:K-ce	8:Bis	9:Em	10:Tok	11:Planeta	12:Mega	13:Karolina	14:Jedynka
1:RMF	.000	2.237	.488	3.206	3.078	4.031	3.042	4.358	3.272	3.201	4.020	3.806	4.516	6.983
2:Blue	2.237	.000	2.255	1.446	2.445	3.540	2.137	2.914	1.981	2.881	3.161	2.136	2.990	5.382
3:ZET	.488	2.255	.000	3.226	3.216	4.119	3.062	4.475	3.303	3.313	4.253	3.718	4.508	7.075
4:Plus	3.206	1.446	3.226	.000	3.511	2.700	1.490	2.135	1.030	2.388	3.863	2.403	2.881	4.405
5:Eska	3.078	2.445	3.216	3.511	.000	5.544	4.291	4.188	3.958	4.897	1.774	3.043	4.299	7.001
6:Trójka	4.031	3.540	4.119	2.700	5.544	.000	2.101	3.042	2.304	1.913	5.529	4.933	4.422	4.294
7:K-ce	3.042	2.137	3.062	1.490	4.291	2.101	.000	2.734	1.215	1.788	4.780	3.417	3.356	4.816
8:Bis	4.358	2.914	4.475	2.135	4.188	3.042	2.734	.000	2.165	3.604	4.035	3.541	3.804	3.947
9:Em	3.272	1.981	3.303	1.030	3.958	2.304	1.215	2.165	.000	2.294	4.253	3.042	2.789	4.416
10:Tok	3.201	2.881	3.313	2.388	4.897	1.913	1.788	3.604	2.294	.000	5.133	4.405	4.194	4.980
11:Planeta	4.020	3.161	4.253	3.863	1.774	5.529	4.780	4.035	4.253	5.133	.000	4.027	4.349	6.347
12:Mega	3.806	2.136	3.718	2.403	3.043	4.933	3.417	3.541	3.042	4.405	4.027	.000	3.483	5.874
13:Karolina	4.516	2.990	4.508	2.881	4.299	4.422	3.356	3.804	2.789	4.194	4.349	3.483	.000	4.292
14:Jedynka	6.983	5.382	7.075	4.405	7.001	4.294	4.816	3.947	4.416	4.980	6.347	5.874	4.292	.000

To jest macierz niepodobierstwa

Źródło: opracowanie własne za pomocą programu SPSS.

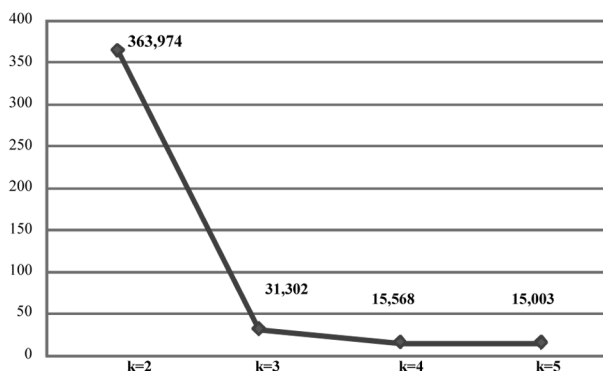
Macierz odległości pomiędzy obiektami (tab. 3) będzie stanowiła podstawę do przeprowadzenia metody iteracyjno- optymalizacyjnej k -średnich oraz skalowania wielowymiarowego.

3. Metoda k -średnich

Algorytmy optymalizacyjne, mające charakter iteracyjny, bazują na założeniu, że znany jest wstępny podział zbioru n obiektów na u klas. Zadaniem tych metod jest poprawienie, z punktu widzenia zdefiniowanej funkcji kryterium, wstępnego podziału zbioru obiektów [Gatnar, Walesiak 2009]. Istota tych metod polega na optymalizowaniu pewnej funkcji jakości podziału obiektów. Chodzi o to, aby zróżnicowanie obiektów w grupach (mierzone za pomocą wariancji) było jak najmniejsze, a pomiędzy grupami jak największe.

Pierwszą metodą klasyfikacji obiektów będzie jedna z metod iteracyjno- optymalizacyjnych, tzn. metoda k -średnich. Liczba klas w metodzie k -średnich zależy od badacza, jednak istnieją kryteria pozwalające na wybór optymalnej liczby klas. Za pomocą indeksu Calińskiego-Harabasa (CH) dokonano sprawdzenia dla $k = 2$, $k = 3$, $k = 4$, $k = 5$ i wynik maksymalny uzyskano dla liczby klas równej 2, a wartość indeksu CH = 363,974.

Wstępne i ostateczne centra skupień dla dwóch klas przedstawia tab. 4. W tym przypadku skupienie 1 jest klasą stacji radiowych o dużej liczbie reklam, z graną muzyką popularną i małą liczbą serwisów kulturalnych, natomiast w skupieniu 2 znajdują się stacje o wysokim poziomie prowadzonych audycji, z graną muzyką dla koneserów i z małą liczbą reklam.

Rys. 1. Wykres zależności indeksu Calińskiego-Harabasa od liczby klas k

Źródło: opracowanie własne.

Tabela 4. Wstępne i ostateczne centra skupień dla $k = 2$ klasy

	Wstępne centra skupień		Ostateczne centra skupień	
	Skupienie		Skupienie	
	1	2	1	2
Reklama	4,35	1,83	4,13	3,17
Antena	2,96	3,83	3,58	3,65
Wiek	2,15	5,17	3,55	4,56
Muzyka	5,81	2,00	4,87	3,12
Informacje	4,62	3,83	3,75	3,74
Polityka	5,21	5,17	4,24	4,69
Poziom	3,98	2,50	3,70	3,19
Promocja	3,15	6,67	3,67	5,00
Pora	3,1	2,2	3,3	2,6
Kultura	5,33	3,83	4,41	3,69

Źródło: opracowanie własne za pomocą programu SPSS.

Przynależność poszczególnych obiektów do skupień oraz odległości obiektów od środków ciężkości w danej klasie przedstawia tab. 5.

Do klasy pierwszej należą 10 obiektów: RMF, Blue, Zet, Plus, Eska, Katowice, Em, Tok, Planeta i Mega. Do klasy drugiej należą cztery obiekty: „Trójka”, Bis, Karolina i „Jedynka”.

Za pomocą analizy wariancji można określić, które zmienne mają największy wpływ na przynależność do danego skupienia, tzn. mają odpowiednią moc dyskryminacyjną. Zmienna, dla której stosunek wariancji międzygrupowej do wariancji wewnątrzgrupowej jest największy, ma znaczny wpływ na przynależność do skupie-

Tabela 5. Przynależność obiektów do skupień dla $k = 2$ klasy

Numer obserwacji	Obiekt	Skupienie	Odległość
1	RMF	1	1,995
2	Blue	1	,588
3	ZET	1	2,078
4	Plus	1	1,369
5	Eska	1	2,501
6	Trójka	2	2,414
7	K-ce	1	1,923
8	Bis	2	1,960
9	Em	1	1,741
10	Tok	1	2,596
11	Planeta	1	3,158
12	Mega	1	2,417
13	Karolina	2	2,670
14	Jedynka	2	2,670

Źródło: opracowanie własne za pomocą programu SPSS.

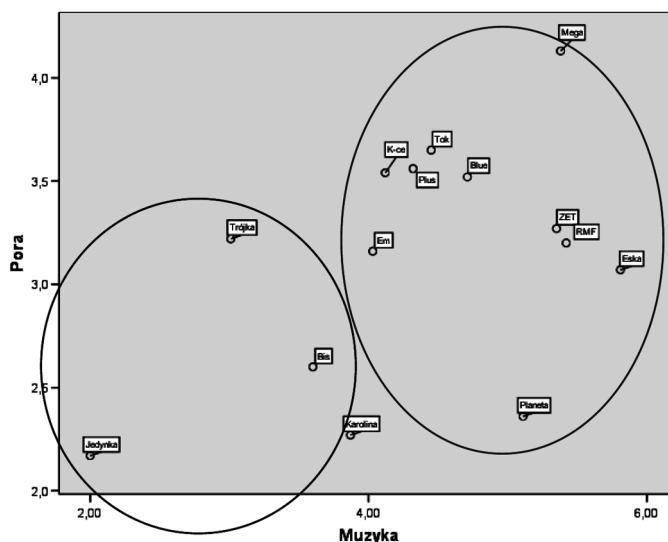
Tabela 6. Analiza wariancji i wartość statystyki F dla poszczególnych zmiennych

	Skupienie		Błąd		F	Istotność
	Średni kwadrat	df	Średni kwadrat	df		
Reklama	2,628	1	,552	12	4,761	,050
Antena	,014	1	,203	12	,067	,800
Wiek	2,906	1	1,032	12	2,816	,119
Muzyka	8,750	1	,464	12	18,864	,001
Informacje	,000	1	,461	12	,001	,977
Polityka	,572	1	1,006	12	,569	,465
Poziom	,736	1	,383	12	1,920	,191
Promocja	5,096	1	,725	12	7,029	,021
Pora	1,743	1	,217	12	8,031	,015
Kultura	1,467	1	,788	12	1,861	,198

Źródło: opracowanie własne za pomocą programu SPSS.

nia. Decydujący wpływ na to, do którego skupienia należy dana stacja radiowa, mają dwie istotne zmienne: rodzaj muzyki ($F = 18,864$) oraz pora ($F = 8,031$).

Przynależność do klas przedstawić można na wykresie rozrzutu ze względu na dwie zmienne dyskryminacyjne, tzn. rodzaj emitowanej muzyki oraz porę nadawania audycji.



Rys. 2. Wykres rozrzutu dla zmiennych muzyka (oś X) oraz pora (oś Y)

Źródło: opracowanie własne za pomocą programu SPSS.

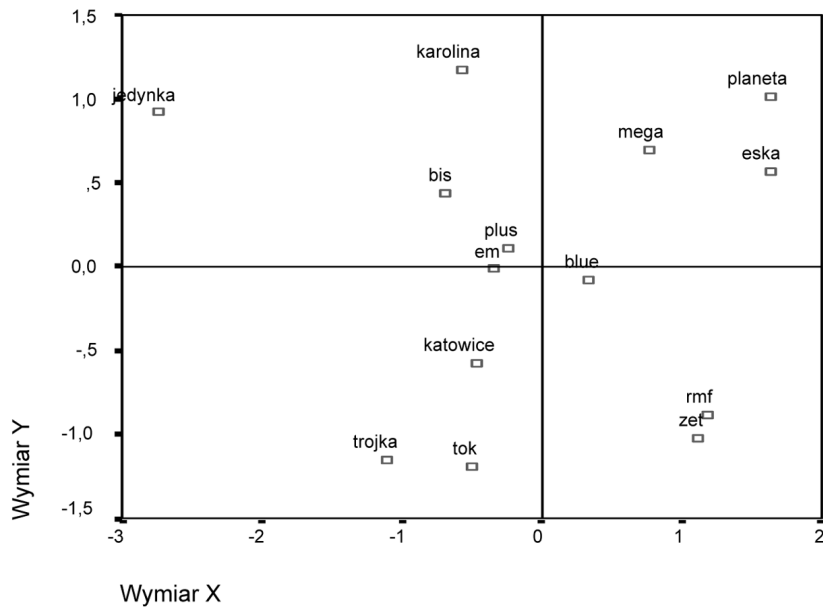
Z wykresu rozrzutu widać, jak przedstawia się rozmieszczenie wszystkich stacji radiowych w przestrzeni dwuwymiarowej. W wyniku podziału metodą k -średnich otrzymano zbiór dwóch rozłącznych klas, z czego klasa pierwsza ma charakter komercyjny, natomiast klasa druga to stacje o profilu informacyjnym.

4. Skalowanie wielowymiarowe

Wyniki klasyfikacji, które otrzymano po przeprowadzeniu metody k -średnich, pozwoliły na segmentację i podział rynku stacji radiowych na jednorodne grupy. Dla weryfikacji otrzymanych wyników przeprowadzono skalowanie wielowymiarowe, którego celem jest ukazanie struktury badanych obiektów przez określenie treści wymiarów na podstawie podobieństw i preferencji respondentów oraz przedstawienie w przestrzeni r -wymiarowej ($r < m$) relacji zachodzących pomiędzy badanymi obiektami. W wyniku zastosowania skalowania wielowymiarowego powstała mapa percepcji, na której ukazano relacje pomiędzy badanymi stacjami radiowymi.

W wyniku zastosowania procedury ALSCAL dla dwóch wymiarów ($r = 2$) otrzymano mapę percepcji, przedstawiającą relacje zachodzące między obiektami (rys. 3).

W celu interpretacji wymiarów powstałych w wyniku zastosowania metody skalowania wielowymiarowego obliczono współczynniki korelacji pomiędzy współrzednymi wymiaru X oraz wymiaru Y a wszystkimi zmiennymi pierwotnymi. Z wymiarem X najsilniej skorelowaną zmienną jest rodzaj muzyki ($r = 0,974$) oraz zmienne: promocja i reklama stacji ($r = -0,849$), zatem wymiar ten jest wymiarem



Rys. 3. Mapa percepcji

Źródło: opracowanie własne za pomocą programu SPSS.

komercyjnym. Z wymiarem *Y* najsilniej skorelowana jest zmienna liczba informacji politycznych ($r = 0,989$) oraz zmienna częstotliwość informacji ($r = 0,899$), co oznacza, że wymiar ten jest wymiarem informacyjnym. Stacje radiowe, które w swoich hasłach reklamowych kładą nacisk na muzykę, to: RMF, Blue, Zet, Eska, Planeta i Mega. Metoda *k*-średnich zaklasyfikowała je do jednej klasy o podobnym profilu radiowym, co oznacza, że słuchacze ocenili te stacje jako zgodne z prezentowanym przez stację profilem. Stacje radiowe, które prezentują się jako stacje informacyjne, to: radiowe „Trójka” i „Jedyńska”, co w metodzie *k*-średnich potwierdza zaklasyfikowanie tych stacji do jednej klasy o podobnej ofercie programowej.

5. Wnioski

W artykule zaprezentowano wyniki jednej z najciekawszych metod analizy danych, tzn. metody *k*-średnich. Do weryfikacji i porównania otrzymanych wyników wykorzystano skalowanie wielowymiarowe. Niniejsze badanie dotyczące rynku radiowego na Śląsku pozwoliło dokonać klasyfikacji stacji radiowych na rozłączne segmenty o podobnych profilach. Wyniki obu metod ze względu na graficzną prezentację stanowią wygodne i użyteczne narzędzie analizy danych, szczególnie w badaniach marketingowych.

W metodzie k -średnich uzyskano 2 rozłączne klasy, natomiast po przeprowadzeniu procedury skalowania wielowymiarowego z wykorzystaniem algorytmu ALSCAL otrzymano graficzne rozmieszczenie obiektów w dwóch wymiarach: wymiarze komercyjnym i informacyjnym.

Stacje radiowe, które w swoich hasłach reklamowych kładą nacisk na muzykę, to: radio RMF, Blue, Zet, Eska, Planeta i Mega. Według metody k -średnich zostały one zaklasyfikowane do jednej klasy, co oznacza, że słuchacze ocenili te stacje jako zgodne z prezentowanym przez stację profilem. Stacje radiowe, które prezentują się jako stacje informacyjne, to: radiowe „Trójka” i „Jedynka”, co metoda k -średnich potwierdza, przydzielając te stacje do jednej klasy.

Literatura

- Borg I., Groenen P., *Modern Multidimensional Scaling, Theory and Applications*, Springer-Verlag, New York 1997.
- Gatnar E., Walesiak M. (red.), *Statystyczna analiza wielowymiarowa z wykorzystaniem programu R*, PWN, Warszawa 2009.
- Gordon A.D., *Classification*, Chapman and Hall, London 1999.
- Green P.E., Rao V.R., *Applied Multidimensional Scaling*, Holt, Rinehart and Winston, New York 1972.
- Hair J.F., Anderson R.E., Tatham R.L., Black W.C., *Multivariate Data Analysis with Readings*, Prentice-Hall, Englewood Cliffs 1995.
- Jajuga K., *Statystyczna analiza wielowymiarowa*, PWN, Warszawa 1993.
- Kendall M.G., *Multivariate Analysis*, Griffin, London 1975.
- Kruskal J.B., *Multidimensional Scaling by Optimising Goodness of Fit to a Nonmetric Hypothesis*, „Psychometrika” 1964 nr 29.
- Mynarski S., *Metody badań marketingowych*, PWE, Warszawa 1990.
- Torgenson W.S., *Multidimensional Scaling: Theory and Method*, „Psychometrika” 1952 nr 17.
- Walesiak M., *Metody analizy danych marketingowych*, PWN, Warszawa 1996.
- Walesiak M., *Statystyczna analiza wielowymiarowa w badaniach marketingowych*, AE, Wrocław 1993.
- Wilkes R.E., *Product Positioning by Multidimensional Scaling*, „Journal of Marketing Research”, sierpień 1977.
- Young F.W., *Multidimensional Scaling*, Wiley, New York 1985.
- Zaborski A., *Skalowanie wielowymiarowe*, AE, Wrocław 2001.

ANALYSIS AND VISUALIZATION OF RADIO MARKET IN SILESIA

Summary

The paper shows two methods of classification together with its graphical results. The methods are used to present different ways of presentation of classification results with the example based on radio market in Silesia. There are two methods shown in this paper: k -means and multidimensional scaling with ALSCAL algorithm. Obtained results show that it is possible to make a market segmentation and present its graphical result.