

Agnieszka Stanimir

Uniwersytet Ekonomiczny we Wrocławiu

ZASTOSOWANIE RÓŻNYCH TECHNIK ANALIZY KORESPONDENCJI W ANALIZIE DANYCH PERSONALNYCH

Streszczenie: Celem artykułu jest wskazanie możliwości zastosowania różnych technik analizy korespondencji w analizie cech personalnych na podstawie *curriculum vitae*. W wielu badaniach marketingowych niezbędne jest jednoczesne poznanie wzajemnych relacji, zależności czy współwystąpień wielu zmiennych. Do tego celu można posłużyć się analizą korespondencji – metodą statystyczną z grupy metod eksploracyjnych. Umożliwia ona zbadanie współwystąpień kategorii cech nominalnych (dwóch lub wielu). Efektem jest graficzna prezentacja wzajemnego położenia kategorii analizowanych zmiennych z uwzględnieniem ich współwystąpień oraz zachowaniem jak największej rzeczywistej informacji o cechach. Do analizy danych wykorzystano macierz Burta, wielowymiarową tablicę kontyngencji, łączoną tablicę kontyngencji oraz własną propozycję połączenia dwóch ostatnich technik.

Słowa kluczowe: analiza danych personalnych, analiza korespondencji, wielowymiarowa tablica kontyngencji, łączona tablica kontyngencji.

1. Wstęp

Celem artykułu jest wskazanie możliwości zastosowania różnych technik analizy korespondencji w analizie cech personalnych na podstawie *curriculum vitae*.

Jednym z etapów badania marketingowego jest analiza zgromadzonych danych, którą wykonuje się niezależnie od tego, czy przeprowadzono badanie ilościowe czy jakościowe. Problemem staje się sposób analizy, a raczej decyzja o zastosowaniu jednej bądź kilku metod statystycznych. W badaniach jakościowych analizy statystyczne wykonywane są sporadycznie, ale nie ma przeciwwskazań do ich stosowania. Podjęcie decyzji o zastosowaniu metod statystycznych i z dziedzin pokrewnych jest uzależnione od poprawnego rozpoznania skali, na której dokonano pomiaru danych (zmiennych). Jest to bardzo ważne ze względu na fakt, że dla danych zmierzonych na skalach słabych (nominalnej, porządkowej) niemożliwe jest zastosowanie technik wykorzystywanych w analizie danych zmierzonych na skalach metrycznych (przedziałowej i ilorazowej). Znane jest natomiast zjawisko transformacji skal, czyli przekształcania pomiarów wykonanych na skalach mocniejszych na pomiary wyko-

nane na skalach słabszych¹. Do analizy pomiarów wykonanych na skali ilorazowej dopuszczalnych jest najwięcej metod analizy statystycznej, dzięki którym można bardzo wnikliwie rozpoznać problem badawczy. Niejednokrotnie uznaje się, że dane zmierzone na skali nominalnej niosą najmniej informacji o badanym zjawisku. Liczba dostępnych metod analizy dla tych danych jest również niewielka. W wielu badaniach marketingowych, poświęcanych na przykład analizie otoczenia przedsiębiorstwa, niezbędne jest jednoczesne poznanie wzajemnych relacji, zależności czy współwystąpień wielu zmiennych. Analiza korespondencji pozwala rozwiązać przedstawione problemy.

Analiza korespondencji należy do eksploracyjnych metod statystycznych. Wykonuje się ją w celu zbadania współwystąpień kategorii cech nominalnych (dwóch lub wielu). Efektem jest graficzna prezentacja w przestrzeni o niskim wymiarze (najczęściej R^2 lub R^3) wzajemnego położenia kategorii analizowanych zmiennych z uwzględnieniem ich współwystąpień oraz zachowaniem jak największej rzeczywistej informacji o cechach. W tym celu przeprowadza się dla odpowiedniej macierzy rozkład według wartości osobliwych, co umożliwi wyznaczenie współrzędnych rzutowania oraz dokonanie oceny jakości odwzorowania. Analizę korespondencji można przeprowadzić dla dwóch zmiennych nominalnych (klasyczna analiza korespondencji) lub większej liczby cech zmierzonych na skali najslabszej (analiza korespondencji wielu zmiennych). W drugim z wymienionych podejść w trakcie analizy wykorzystuje się macierz Burta, łączoną tablicę kontyngencji lub wielowymiarową tablicę kontyngencji. Analizę korespondencji można też stosować w przypadku danych porządkowych².

W przeprowadzonych analizach wykorzystano dane pochodzące ze zgłoszeń kandydatów ubiegających się o posadę programisty w Firmie X, tworzącej wysoko wyspecjalizowane programy komputerowe. Od kandydatów na stanowisko programisty wymagano: wykształcenia wyższego kierunkowego, znajomości oprogramowania, znajomości języka angielskiego. Przeanalizowano dane z CV i listów motywacyjnych 126 kandydatów³. Dogłębna analiza posiadanych danych jest nie-

¹ Zob. J. Steczkowski, A. Zeliaś, *Metody statystyczne w badaniach cech jakościowych*, Wydawnictwo AE w Krakowie, Kraków 1997; M. Walesiak, *Metody analizy danych marketingowych*, PWN, Warszawa 1996.

² Techniczne aspekty prowadzenia analizy korespondencji we wszystkich podejściach opisano w pracach: A. Stanimir, *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo AE we Wrocławiu, Wrocław 2005; *Analiza danych marketingowych. Problemy, metody, przykłady*, Wydawnictwo AE we Wrocławiu, Wrocław 2006.

³ Zob. analizę rynku pracy na podanym stanowisku w: A. Stanimir, A. Grześkowiak, *Wykorzystanie wielowymiarowych technik analizy danych w ocenie umiejętności osób ubiegających się o pracę*, referat wygłoszony na XIV Konferencji Naukowej w cyklu Warsztaty Metodologiczne z Metod Ilościowych w Badaniach Marketingowych im. Profesora Stefana Mynarskiego pt. „Wyzwania badawcze i analityczne współczesnego marketingu”, 31 maja – 1 czerwca 2010 r., Szkoła Główna Handlowa w Warszawie, w druku.

wątpliwie uzupełnieniem analizy rynku pracy, którą Firma prowadzi od początku swojej działalności.

Z listów motywacyjnych i CV kandydatów wyodrębniono następujące cechy demograficzne oraz charakteryzujące umiejętności:

- płeć (1 – kobieta; 0 – mężczyzna);
- wiek (do 25 lat – A; 25-29 – B; 30 i więcej – C);
- miejsce zamieszkania (1 – Wrocław; 2 – poza Wrocławiem);
- doświadczenie zawodowe (1 – brak; 2 – pracował/a w 1-2 miejscach; 3 – pracował/a w wielu miejscach);
- zgodność doświadczenia zawodowego ze stanowiskiem, na które aplikuje (1 – tak; 0 – nie);
- wykształcenie (1 – średnie; 2 – policealne; 3 – wyższe);
- znajomość języka angielskiego (1 – podstawowa; 2 – dobra; 3 – biegła).

Cechy: wykształcenie, wiek i znajomość języka angielskiego są mierzone na skali porządkowej, a pozostałe cechy na skali nominalnej. Przy każdej zmiennej podano w nawiasach jej kategorie z przyjętym sposobem kodowania.

2. Zastosowanie macierzy Burta w analizie umiejętności kandydatów

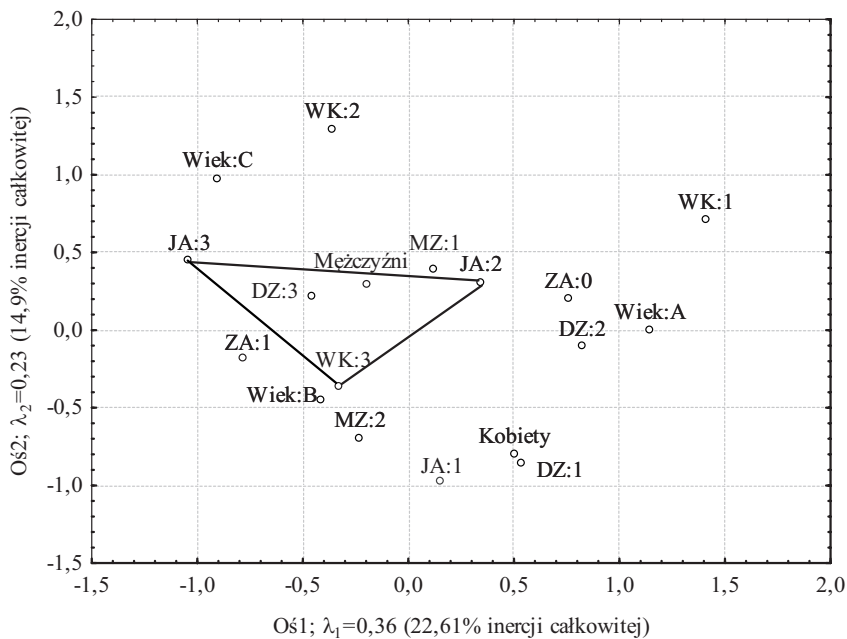
Analiza korespondencji wielu cech jest najczęściej prowadzona z wykorzystaniem macierzy Burta. Macierz ta jest symetryczna i ma budowę blokową. Na głównej przekątnej znajdują się macierze diagonalne, zawierające liczebności wystąpień każdej kategorii analizowanych cech. Liczba macierzy jest zgodna z liczbą zmiennych. Poza przekątną umieszczone są tablice kontyngencji dla każdej pary cech (tab. 1).

Na rysunku 1 zamieszczono podstawowe informacje o jakości odwzorowania współwystąpień w przestrzeni R^2 . Dwie pierwsze wartości własne λ_1 oraz λ_2 , na podstawie których tworzona jest przestrzeń dwuwymiarowa, stanowią 37% inercji całkowitej. Analizując rozrzut punktów obrazujących poszczególne kategorie cech, należy zwrócić uwagę na odległość punktu od początku układu współrzędnych oraz bliskość wzajemnego położenia punktów. W pierwszym przypadku punkty leżące blisko początku układu współrzędnych wnoszą najmniej informacji o współwystąpieniach kategorii cech, a punkty najbardziej oddalone – najwięcej. Jeśli natomiast punkty obrazujące kategorie opisujące różne cechy są położone blisko siebie, to znaczy, że te kategorie często współwystępują. Na rysunku połączono ze sobą punkty obrazujące požądane przez pracodawcę kategorie cech. Można stwierdzić, że osoby, które znają język angielski w stopniu co najmniej dobrym i mają właściwe wykształcenie, to mężczyźni z dużym doświadczeniem zawodowym (DZ: 3). Wśród kobiet ubiegających się o pracę programisty dominują osoby słabo znające język angielski (JA: 0), bez doświadczenia zawodowego (DZ: 1). Najmłodsze (wiek: A) osoby wśród zgłaszających swoje kandydatury pracowały w jednym lub dwóch miejscach (DZ: 2), ale nie jako programiści (ZA: 0).

Tabela 1. Przykład budowy macierzy Burta na podstawie trzech cech opisujących kandydatów

		Płeć		Wiek			Wykształcenie		
		mężczyzna	kobieta	do 25 lat	25-29	30 i więcej	średnie	policealne	wyższe
Płeć	mężczyzna	93	0	24	24	45	21	9	63
	kobieta	0	33	12	18	12	3	3	27
Wiek	do 25 lat	24	3	27	0	0	0	3	24
	25-29	24	18	0	42	0	18	3	21
	30 i więcej	45	12	0	0	57	6	6	45
Wykształcenie	średnie	21	3	0	18	6	24	0	0
	policealne	9	3	3	3	6	0	12	0
	wyższe	63	27	24	21	45	0	0	90

Źródło: opracowanie własne.



Rys. 1. Graficzna prezentacja współwystępowania kategorii wszystkich cech dla kandydatów na stanowisko programisty (MZ – miejsce zamieszkania; DZ – doświadczenie zawodowe; ZA – zgodność doświadczenia zawodowego z aplikacją; WK – wykształcenie; JA – znajomość języka angielskiego)

Źródło: opracowanie własne.

Z punktu widzenia właściciela firmy interesujące jest zbadanie współwystąpień wszystkich cech, jakie można wyróżnić na podstawie CV. Graficzna prezentacja jednoczesnych wystąpień kategorii analizowanych siedmiu cech nie jest możliwa, gdyż musiałaby to być przestrzeń o wymiarze $K = J - Q = 22 - 8 = 14$ (J – to łączna liczba kategorii wszystkich cech, Q – liczba cech). Analiza korespondencji pozwala na prezentację współwystąpień kategorii tych cech w przestrzeni dwuwymiarowej (zob. rys. 1).

3. Wielowymiarowa i łączona tablica kontyngencji w analizie danych personalnych

Wielowymiarowa tablica kontyngencji jest krzyżową klasyfikacją wielu zmiennych. Jeśli analizę prowadzi się dla trzech cech, to tablicę taką można stworzyć z wierszy, kolumn i warstw. Sposób tworzenia wielowymiarowej tablicy kontyngencji komplikuje się wraz ze wzrostem liczby analizowanych cech. W takim przypadku warstwy mogą wystąpić również w kolumnach. W tabeli 2 zaprezentowano wielowymiarową tablicę kontyngencji dla czterech wybranych cech: płeć, wiek, wykształcenie, znajomość języka angielskiego.

Tabela 2. Wielowymiarowa tablica kontyngencji cech: płeć, wiek, wykształcenie, znajomość języka angielskiego

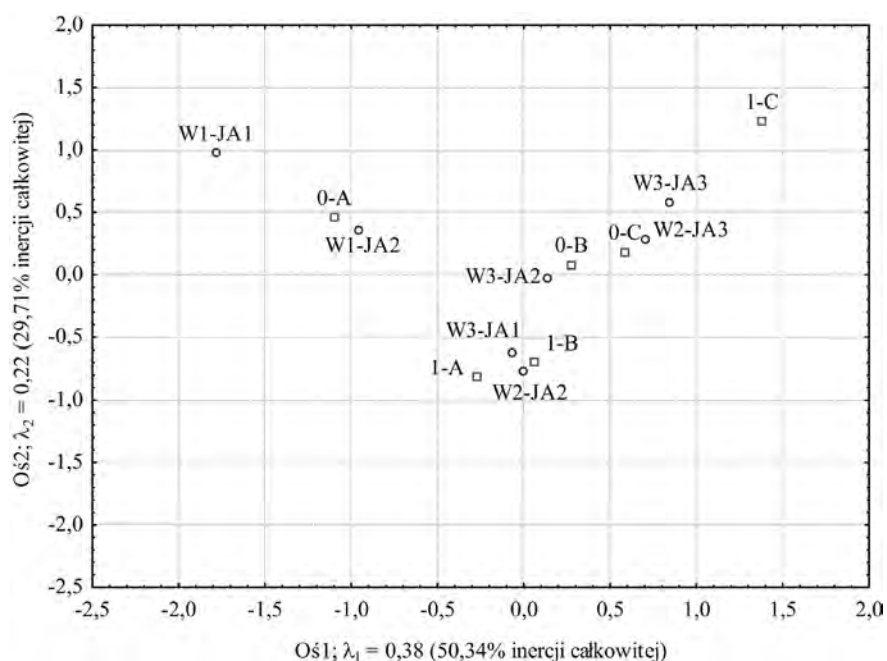
		Płeć + wiek = PW					
		0-A	0-B	0-C	1-A	1-B	1-C
Wykształcenie + język angielski = WKJA	WK1-JA1	3	0	0	0	0	0
	WK1-JA2	12	6	0	3	0	0
	WK1-JA3	0	0	0	0	0	0
	WK2-JA1	0	0	0	0	0	0
	WK2-JA2	0	3	0	3	0	0
	WK2-JA3	0	3	3	0	0	0
	WK3-JA1	3	9	3	9	6	0
	WK3-JA2	6	12	12	3	6	0
	WK3-JA3	0	12	6	0	0	3

Płeć: 0 – mężczyzna, 1 – kobieta; wiek: A – do 25 lat; B – 25-29 lat, C – 30+; wykształcenie: WK1 – średnie, WK2 – policealne, WK3 – wyższe; znajomość języka angielskiego: JA1 – podstawowa, JA2 – dobra, JA3 – biegła

Źródło: opracowanie własne.

Kategorie cech kombinowanych PW oraz WKJA powstają przez połączenie kategorii składowych cech, odpowiednio: płeć i wiek oraz wykształcenie i znajomość

mość języka angielskiego. Z udziałem tych czterech cech pierwotnych można stworzyć również inne tablice wielowymiarowe np.: [PWK][WJA]; [WWK][PJA]. Jeśli w prowadzonym badaniu nie ma wskazania, jakie kombinacje zmiennych są istotne dla rozwiązania problemu badawczego, to utworzone dla różnych kombinacji cech tablice wielowymiarowe są poddawane ocenie za pomocą testu niezależności χ^2 lub współczynnika największej wiarygodności⁴. W celu stwierdzenia, jak silna jest zależność cech, można obliczyć jeden ze wskaźników: ϕ^2 , V -Cramera, C -Pearsona. Nie jest zalecane stosowanie współczynnika T -Czuprowa ze względu na duże rozmiary tablic. Po dokonaniu wyboru właściwej tabeli przeprowadza się analizę korespondencji zgodnie z podejściem klasycznym. Wymiar rzeczywistej przestrzeni rzutowania jest równy $\min(r-1; c-1)$, gdzie r to liczba wierszy (kategorii cechy zapisanej w wierszach), c to liczba kolumn (kategorii cechy zapisanej w kolumnach). Budując wielowymiarowe tablice kontyngencji, należy pamiętać, by sumy brzegowe wierszy oraz kolumn były większe od zera. W przypadku przeciwnym nie można przeprowadzić analizy korespondencji.



Rys. 2. Graficzna prezentacja współwystępowania kategorii cech [PW] oraz [WKJA] na podstawie analizy korespondencji wielowymiarowej tablicy kontyngencji (oznaczenia jak w tab. 2)

Źródło: opracowanie własne.

⁴ S.E. Clausen, *Applied Correspondence Analysis. An Introduction*, University Paper 121, Sage, 1998; P.G.M. van der Heijden, *Correspondence Analysis of Longitudinal Categorical Data*, DSWO Press, Leiden 1987.

Na podstawie danych zawartych w tabeli 2 nie można przystąpić do przeprowadzenia analizy korespondencji, gdyż dla dwóch kategorii cechy kombinowanej [WKJA] sumy brzegowe są równe zero. W dalszej analizie zrezygnowano więc z kategorii opisujących osoby z wykształceniem średnim i biegłą znajomością języka angielskiego oraz z wykształceniem policealnym i podstawową znajomością języka angielskiego. Wynik zaprezentowano na rysunku 2.

Prezentacja współwystąpień kategorii cech [PW][WKJA] w przestrzeni dwuwymiarowej zachowuje 80% rzeczywistych relacji (50,34% + 29,71%). Rzeczywista przestrzeń rzutowania jest pięciowymiarowa. Analizując rozrzut punktów na wykresie, należy stwierdzić, że osoby z najlepszą znajomością języka angielskiego, wykształceniem policealnym lub wyższym to mężczyźni mający ponad 29 lat.

Innym podejściem w analizie korespondencji jest technika wykorzystująca zasadę addytywności statystyki chi-kwadrat⁵. Reguła ta znajduje zastosowanie w zbiorach danych opisujących zbliżone pod względem wielkości i struktury doświadczenia. Statystyki chi-kwadrat oraz stopnie swobody wyznaczone dla każdego zbioru można zsumować i wyciągać wnioski o całym rozpatrywanym zjawisku. Działanie to jest bardzo użyteczne, gdy w trakcie badania niezbędne okazuje się zanalizowanie współwystąpień kategorii jednej cechy (cechy wspólnej) z kategoriami pozostałych cech. Możliwe w takim przypadku jest kilkakrotne zastosowanie klasycznej analizy korespondencji, ale działanie to jest czasochłonne, a prezentacja wyników zbyt obszerna. Wyodrębniwszy we wszystkich pomiarach cechę wspólną, należy następnie zbudować łączoną tablicę kontyngencji. Przykład takiej tablicy zapisano w tabeli 3.

Dla danych zapisanych w tabeli 3 przeprowadzono analizę korespondencji zgodnie z podejściem klasycznym. Wynik prezentuje rysunek 3. Rzeczywista przestrzeń powiązań kategorii cech jest dwuwymiarowa ($\min(r-1; c-1)$), zatem prezentacja w dwuwymiarowym układzie współrzędnych wyjaśnia 100% inercji całkowitej (81,04% + 18,96%).

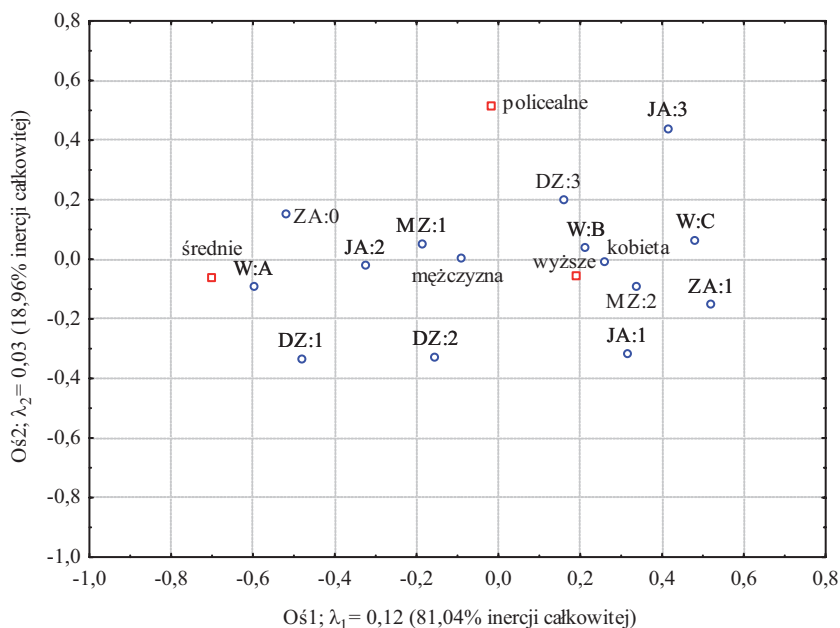
Analizując rozrzut punktów na rysunku 3, należy pamiętać, by nie wskazywać współwystąpień kategorii między poszczególnymi cechami, lecz jedynie łączne wystąpienia kategorii cechy wspólnej z kategoriami pozostałych cech. Punkt obrazujący osoby z wykształceniem wyższym jest położony blisko punktu prezentującego kobiety. Wykształcenie wyższe jest charakterystyczne dla osób w wieku 25-29 lat. Jednak nie można powiedzieć, że wykształcenie wyższe mają kobiety w wieku 25-29 lat. Osoby z wykształceniem średnim to te, które należą do najmłodszych w badaniu. Osoby o takim poziomie wykształcenia najczęściej nie mają doświadczenia zawodowego. Punkt obrazujący osoby z wykształceniem policealnym jest położony z dala od punktów obrazujących kategorie pozostałych cech, zatem nie można dla nich określić jednoczesnych wystąpień z kategoriami pozostałych cech.

⁵ Reguła ta jest szczegółowo omówiona w pracy: G.U. Yule, M.G. Kendall, *Wstęp do teorii statystyki*, PWN, Warszawa 1966.

Tabela 3. Łączona tablica kontyngencji z cechą wspólną: wykształcenie

		Wykształcenie		
		1	2	3
Płeć	0	21	9	63
	1	3	3	27
Wiek	A	18	3	21
	B	6	6	45
	C	0	3	24
Miejsce zamieszkania	1	21	9	51
	2	3	3	39
Doświadczenie zawodowe	1	6	0	9
	2	9	0	24
	3	9	12	57
Doświadczenie zgodne ze stanowiskiem	0	24	9	30
	1	0	3	60
Znajomość języka angielskiego	1	3	0	30
	2	21	6	39
	3	0	6	21

Źródło: opracowanie własne.

**Rys. 3.** Graficzna prezentacja współwystępowania kategorii cech na podstawie analizy korespondencji łączonej tablicy kontyngencji z cechą wspólną: wykształcenie

Źródło: opracowanie własne.

4. Propozycja kombinacji dwóch podejść w analizie korespondencji

W badaniach marketingowych często analizowane są reakcje różnych zjawisk na wystąpienie pewnych czynników. Podobnie w przypadku badania danych personalnych. Z punktu widzenia pracodawcy istotne jest, by kandydaci na stanowisko programisty mieli wyższe wykształcenie kierunkowe oraz znali język angielski. W celu właściwego ocenienia popytu na pracę na tym stanowisku oraz uzyskania informacji, którą będzie można wykorzystać w przyszłości, istotne staje się zbadanie współwystąpień kategorii pożądaných cech z kategoriami pozostałych zmiennych. W takiej sytuacji autorka proponuje zastosowanie kombinacji dwóch podejść w analizie korespondencji: utworzenie cechy kombinowanej z cech: wykształcenie i znajomość języka angielskiego, uznanie jej dalej jako cechy wspólnej w łączonej tablicy kontyngencji, a następnie przeprowadzenie analizy korespondencji badającej współwystąpienia kategorii cechy wspólnej z kategoriami pozostałych cech odrębnie.

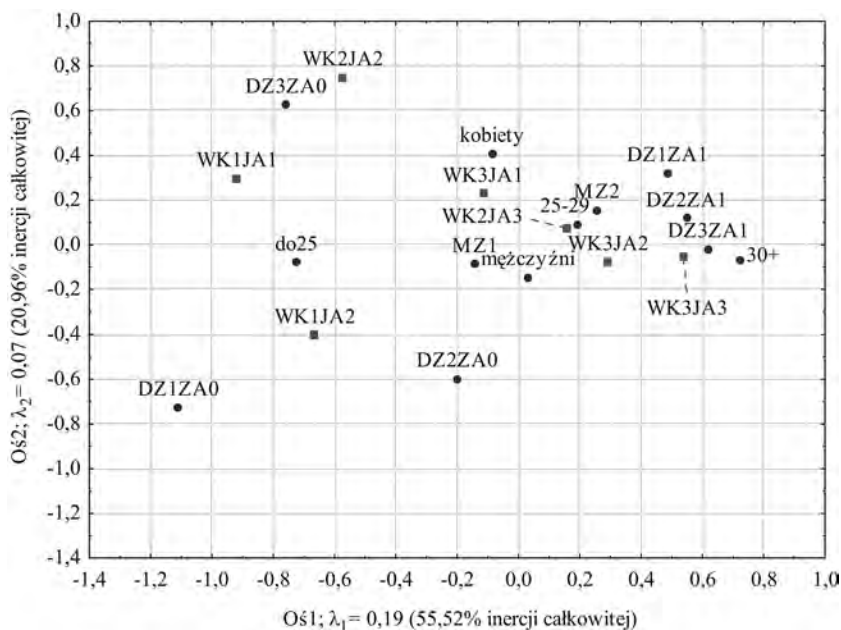
Postępując w zaproponowany sposób, stworzono cechę kombinowaną, zawierającą kategorie: WK1JA1 (wykształcenie średnie + podstawowa znajomość języka angielskiego), WK1JA2 (wykształcenie średnie + dobra znajomość języka angielskiego), WK1JA3 (wykształcenie średnie + biegła znajomość języka angielskiego), WK2JA1 (wykształcenie policealne + podstawowa znajomość języka angielskiego), WK2JA2, WK2JA3, WK3JA1, WK3JA2, WK3JA3. Powstała zmienna i jej kategorie uznano dalej za cechę wspólną i przeprowadzono analizę korespondencji łączonej tablicy kontyngencji, uwzględniając pięć pozostałych cech. Dodatkowo stworzono jeszcze jedną cechę kombinowaną z: doświadczenie zawodowe i zgodność doświadczenia z aplikowanym stanowiskiem. Schemat tablicy na podstawie, której przeprowadzono analizę, zawiera tabela 4.

W tabeli 4 kolorami zaznaczono poszczególne tablice kontyngencji, w których jedna z cech zawsze jest taka sama. W badaniu nie wystąpiły osoby, które mają wykształcenie średnie i bardzo dobrze oceniały swoją znajomość języka angielskiego, oraz te, które mają wykształcenie policealne i słabo znają język angielski. Te dwie kategorie odrzucono. Na podstawie zaprezentowanej tabeli przeprowadzono klasyczną analizę korespondencji, a wyniki przedstawiono na rysunku 4. Rzeczywista przestrzeń powiązań kategorii cech jest sześciowymiarowa ($\min(r-1; c-1)$). Prezentacja w przestrzeni R^2 łączonej tablicy kontyngencji z kombinowaną cechą wspólną odzwierciedla 76% rzeczywistych współwystąpień kategorii zmiennych. Punkty obrazujące wyższe wykształcenie i dobrą lub biegłą znajomość języka angielskiego (WK3JA2, WK3JA3) są położone blisko siebie. Osoby z wyższym wykształceniem i biegłą znajomością języka angielskiego mają ponad 29 lat. Osoby te najczęściej mają doświadczenie zawodowe zgodne ze stanowiskiem, na które aplikują. Osoby z wyższym wykształceniem, ale słabszą znajomością języka są młodsze. Doświadczenie zawodowe tych osób też jest zgodne z posadą, o jaką się ubiegają.

Tabela 4. Łączona tablica kontyngencji ze wspólną cechą kombinowaną

		Wykształcenie + znajomość języka angielskiego						
		WK1JA1	WK1JA2	WK2JA2	WK2JA3	WK3JA1	WK3JA2	WK3JA3
Płeć	mężczyźni							
	kobiety							
Wiek	do 25							
	25-29							
	30 i więcej							
Miejsce zamieszkania	MZ1							
	MZ2							
Doświadczenie zawodowe + zgodność z aplikacją	DZ1ZA0							
	DZ1ZA1							
	DZ2ZA0							
	DZ2ZA1							
	DZ3ZA0							
	DZ3ZA1							

Źródło: opracowanie własne.

**Rys. 4.** Graficzna prezentacja współwystępowania kategorii cech na podstawie analizy korespondencji łączonej tablicy kontyngencji z cechą wspólną kombinowaną: wykształcenie + Znajomość języka angielskiego

Źródło: opracowanie własne.

5. Podsumowanie

Każda z przedstawionych technik analizy korespondencji niesie inny zakres informacji o badanym zjawisku. W przypadku analizy wykonanej na podstawie macierzy Burta prezentowane są jednocześnie związki między kategoriami wszystkich cech. Można zatem badać wzajemne położenia wszystkich punktów względem siebie. Analiza rozmieszczenia punktów na wykresie w przypadku kolejnych technik jest już inna. W wielowymiarowej analizie korespondencji ogólnie zadane są (wybrane przez badacza) związki między pewnymi cechami. Można więc badać współwystąpienia nowych, połączonych kategorii cech kombinowanych. W łączonej tablicy kontyngencji zakłada się pewną cechę nadrzędną, wspólną wobec pozostałych. Z tego powodu możliwe jest rozpatrywanie współwystąpień kategorii tej cechy z kategoriami pozostałych cech, ale nie jednocześnie. Połączenie dwóch ostatnich technik analizy korespondencji niesie dodatkowe możliwości rozpoznawania współwystąpień. Wyróżnienie jako cechy wspólnej zmiennej kombinowanej umożliwia wprowadzenie do badania powiązań między cechami, które są wyróżnione przez badacza lub zlecniodawcę badania.

Prowadząc analizę korespondencji, należy pamiętać, że sumy brzegowe analizowanych tabel zawierających liczbę wystąpień kategorii cech nie mogą być równe zero. Ponadto tworząc cechy kombinowane, należy pamiętać o zwiększającej się liczbie kategorii, co przyczynia się do występowania kolumn lub wierszy o sumach brzegowych równych zero. W takim przypadku należy w dalszej analizie zrezygnować z takiej kategorii. Duża liczba kategorii cech przy dużej liczbie zmiennych powoduje trudności w trakcie odczytywania etykiet punktów na wykresie.

Zastosowanie analizy korespondencji w badaniu danych personalnych jest postępowaniem wspomagającym podejmowanie decyzji personalnych oraz oceny postrzegania przedsiębiorstwa przez pracobiorców na rynku. W przedstawionych analizach wykorzystano jedynie kilka informacji z dostępnych CV. Zakres analizy można zwiększyć, wprowadzając do badania dodatkowe cechy, np. miejsce urodzenia, ukończone kierunki studiów, miejsca ukończenia nauki, długość nauki, liczba szkoleń itd.

Literatura

- Analiza danych marketingowych. Problemy, metody, przykłady*, Wydawnictwo AE we Wrocławiu, Wrocław 2006.
- Clausen S.E., *Applied Correspondence Analysis. An Introduction*, University Paper 121, Sage, 1998.
- Heijden P.G.M. van der, *Correspondence Analysis of Longitudinal Categorical Data*, DSWO Press, Leiden 1987.
- Stanimir A., *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo AE we Wrocławiu, Wrocław 2005.

Steczkowski J., Zeliaś A., *Metody statystyczne w badaniach cech jakościowych*, Wydawnictwo AE w Krakowie, Kraków 1997.

Walesiak M., *Metody analizy danych marketingowych*, PWN, Warszawa 1996.

Yule G.U., Kendall M.G., *Wstęp do teorii statystyki*, PWN, Warszawa 1966.

THE USE OF DIFFERENT TECHNIQUES OF CORRESPONDENCE IN THE ANALYSIS OF PERSONAL DATA

Summary: The aim of the article is to indicate the opportunity of using different techniques of correspondence analysis in the study of personal characteristics on the basis of curriculum vitae. In a lot of marketing research (for example devoted to company environment, the perception of the company by employees) it is necessary to present at the same time the relationship, dependences or joint observation of a lot of variables. To achieve this goal a correspondence analysis was proposed. The task of this method is an indication of multidimensional associations among the categories of nominal variables (two or more). The result is a graphical representation of the relationships among the categories of analyzed variables in terms of their joint associations and preserving as much of the real information of variables as possible. For the data study an analysis based on the Burt matrix is proposed, a multiway contingency table, the concatenated contingency table and the author's own proposal of the combination of two previous techniques.