

**Małgorzata Machowska-Szewczyk,
Agnieszka Sompolska-Rzechuła**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

ANALIZA KORESPONDENCJI W BADANIU POSTAW OSÓB DOKONUJĄCYCH ZAKUPÓW PRZEZ INTERNET

Streszczenie: W pracy podjęto temat analizy częstości kupowania przez Internet w środowisku studentów wybranych szczecińskich uczelni. Celem badania była próba określenia determinant wpływających na częstość dokonywania zakupów za pośrednictwem Internetu. Do osiągnięcia celu wykorzystano analizę korespondencji, która pozwoliła na określenie powiązań występujących między kategoriami zmiennych przyjętych do opisu zjawiska kupowania przez Internet, jak np. miejsce dokonywania zakupów, powody dokonywania zakupów, najczęściej kupowane produkty, forma dokonywania płatności oraz transportu. Spośród zmiennych demograficzno-społecznych w badaniu uwzględniono jedynie płeć kupujących, pozostałe zmienne wykazywały niezależność ze zmienną określającą częstość zakupów przez Internet.

Słowa kluczowe: analiza korespondencji, zakupy przez Internet, kwestionariusz ankiety.

1. Wstęp

Zakupy przez Internet stają się coraz bardziej popularne. Ten sposób nabywania towarów pozwala zaoszczędzić czas, a często także pieniądze. Porównanie ofert oraz cen w sklepach wirtualnych jest dużo łatwiejsze i szybsze niż w tradycyjnych. Niewątpliwą wadą tego rodzaju zakupów jest brak możliwości dokładnego obejrzenia towaru. Opisy przedmiotów oferowanych przez sklepy internetowe są często zbyt skąpe, a zdjęcia małe lub niewystarczająco wyraźne. Jak wynika z badań CBOS z czerwca 2009 r. – przybywa kupujących przez Internet. W czerwcu 2009 r. blisko 32% Polaków powyżej osiemnastego roku życia deklarowało, że dokonali kiedykolwiek zakupów przez Internet. Analiza zjawiska e-zakupów w poszczególnych miesiącach pokazuje, że wzrost liczby kupujących przez Internet nie przekłada się na regularność kupowania.

W pracy podjęto temat analizy częstości kupowania przez Internet w środowisku studentów wybranych szczecińskich uczelni. Celem badania była próba określenia determinant wpływających na częstość dokonywania zakupów za pośrednictwem Internetu. Do osiągnięcia celu wykorzystano analizę korespondencji, która pozwoliła także na określenie powiązań występujących między kategoriami zmiennych przyjętych do opisu zjawiska kupowania przez Internet.

2. Opis metody¹

Analiza korespondencji wielu zmiennych jest narzędziem służącym do intuicyjnego wnioskowania o powiązaniach zachodzących między kategoriami cech typu jakościowego. Punktem wyjścia jest zapisanie złożonej tablicy kontyngencji, w której umieszcza się liczebności wystąpień poszczególnych kategorii zmiennych przyjętych do opisu n obiektów. W zagadnieniach praktycznych bardzo często stosowanym sposobem zapisu danych jest macierz Burta. Utworzenie tej macierzy wymaga budowy złożonej macierzy znaczników \mathbf{Z} , która składa się z podmacierzy odpowiadających poszczególnym zmiennym, czyli $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_Q]$, gdzie Q oznacza liczbę cech. Elementy złożonej macierzy znaczników przyjmują wyłącznie wartości 1 lub 0, w zależności od tego, czy dany obiekt posiada wyróżnioną kategorię zmiennej, czy też nie.

Tabela 1. Przykładowa podmacierz \mathbf{Z} , macierzy znaczników

Numer obiektu	Kategorie cechy numer i			
	1_i	2_i	...	J_i
1	0	1	...	0
2	1	0	...	0
3	0	0	...	1
...
n	0	1	...	0

Źródło: opracowanie własne na podstawie [Stanimir 2005].

Macierz Burta powstaje w wyniku działania $\mathbf{B} = \mathbf{Z}^T \mathbf{Z} = [b_{ij}]_{J \times J}$, $J = \sum_{q=1}^Q J_q$, gdzie J_q jest liczbą kategorii cechy q . Macierz ta jest symetryczna, poza główną przekątną znajdują się tablice kontyngencji odpowiadające dwóm różnym zmiennym i zawierające liczby obiektów o ustalonych kategoriach tych dwóch cech. Na głównej przekątnej umieszczone są macierze diagonalne, w których wartości niezerowe oznaczają liczebności wystąpień danej kategorii cechy.

¹ Opracowano na podstawie [Stanimir 2005; Stanisiz 2007].

Ważnym pojęciem w analizie korespondencji są profile wierszy oraz kolumn. Na podstawie macierzy Burta można wyznaczyć macierz częstości zaobserwowanych \mathbf{P} , dzieląc liczebności w poszczególnych komórkach przez sumę liczebności wszystkich jej cel, czyli nQ^2 . Macierz profili wierszowych otrzymuje się, dzieląc częstości względne w każdym wierszu macierzy \mathbf{P} przez sumę wszystkich częstości w danym wierszu. Analogicznie macierz profili kolumnowych można wyznaczyć, dzieląc każdą częstość względną w kolumnie macierzy \mathbf{P} przez sumę częstości w danej kolumnie. Przeciętny profil wierszowy (kolumnowy) można uzyskać przez sumowanie częstości względnych w każdej kolumnie (wierszu).

Dowolny wiersz oraz kolumna macierzy profili może zostać przedstawiona jako punkt w przestrzeni, ale wymiary tych przestrzeni zwykle nie pozwalają na graficzną prezentację.

Głównym celem analizy korespondencji jest przedstawienie analizowanego zbioru danych w przestrzeni maksymalnie trójwymiarowej, przy zachowaniu możliwie najpełniejszej informacji o zróżnicowaniu wierszy oraz kolumn. Do tego celu służy metoda rozkładu macierzy według wartości osobliwych. Zgodnie z podejściem Greenacre'a rzutowanie na przestrzeń o niskim wymiarze odbywa się przez rozkład macierzy \mathbf{A} według wartości osobliwych, czyli:

$$\mathbf{A} = \mathbf{D}_r^{-1/2} (\mathbf{P} - \mathbf{r}\mathbf{r}^T) \mathbf{D}_r^{-1/2} = \mathbf{U}\mathbf{\Gamma}^2\mathbf{U}^T,$$

gdzie: \mathbf{P} – macierz częstości zaobserwowanych $\mathbf{P} = \frac{1}{nQ^2} \mathbf{B}$,

$\mathbf{r} = [p_{i\cdot}]$ – wektor częstości brzegowych wierszy, przy czym $p_{i\cdot} = \frac{Q \cdot b_{ii}}{n \cdot Q^2}$;

\mathbf{D}_r – macierz diagonalna częstości brzegowych wierszy,

\mathbf{U} – macierz wartości własnych $\lambda_{B,k}$ macierzy \mathbf{A} ,

$\mathbf{\Gamma}^2 = [\gamma_k^2]$ – macierz diagonalna zawierająca kwadraty wartości osobliwych

γ_k macierzy \mathbf{A} , $k \in \{1, 2, \dots, K\}$, $K = \sum_{q=1}^Q (J_q - 1)$, J_q – liczba kategorii cechy q .

Wartości osobliwe macierzy \mathbf{A} odpowiadają pierwiastkom kwadratowym wartości własnych, można zatem zapisać: $\gamma_k^2 = \lambda_k$, $k \in \{1, \dots, K\}$. W analizie korespondencji używa się pojęcia bezwładności, które jest analogiczne do wariancji. Całkowita bezwładność jest miarą rozproszenia profili wokół odpowiednich profili przeciętnych. Można wykazać, że całkowita bezwładność wierszy jest równa całkowitej bezwładności kolumn, stąd rozważa się tylko jedną wartość zwaną bezwładnością całkowitą lub inercją całkowitą. Jeżeli inercja całkowita jest bliska zeru, to oznacza niewielką różnicę między profilami a profilem przeciętnym, a tym

samym szansę wystąpienia istotnego powiązania między kategoriami poszczególnych zmiennych. Całkowita inercja jest sumą kwadratów wartości własnych, czyli $\lambda = \sum_{k=1}^K \lambda_k$. Wartości λ_k , $k \in \{1, \dots, K\}$ są nazywane inercjami głównymi. W zagadnieniach praktycznych stosuje się tzw. procent bezwładności, czyli λ_k/λ , oznaczający tę część bezwładności pierwotnej, która została wyjaśniona przez k -ty czynnik.

Poprzez rozkład macierzy \mathbf{A} według wartości osobliwych definiowany jest układ współrzędnych, w który rzutowane będą punkty odpowiadające kolejnym kolumnom (wierszom). Wyznacza się macierz \mathbf{F} , zawierającą współrzędne kategorii na wszystkich osiach powiązań następująco:

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{U} \mathbf{T}_B.$$

Wymiar przestrzeni rzeczywistych powiązań, otrzymanej na podstawie analizy macierzy Burta, wynosi $K = \sum_{q=1}^Q (J_q - 1)$. W dalszej części badania dokonuje się rzutowania na przestrzeń o niższym wymiarze, tak aby strata informacji była jak najmniejsza. Rozmiar przestrzeni rzutowania można określić na podstawie wartości miernika, który wskazuje na udział inercji ustalonego wymiaru w bezwładności całkowitej (procent skumulowany) wyznaczanego zgodnie ze wzorem:

$$\tau_{K^*} = \frac{\sum_{k=1}^{K^*} \lambda_k}{\sum_{k=1}^K \lambda_k} = \frac{\sum_{k=1}^{K^*} \lambda_k}{\lambda},$$

gdzie: $K^* \in \{1, \dots, K\}$ oznacza wybrany wymiar rzutowania.

W sytuacji gdy macierz Burta jest bardzo rozbudowana, rzeczywisty wymiar przestrzeni współwystępowania kategorii jest bardzo wysoki. Pierwsza wartość własna nie jest wtedy znacząco wyższa od pozostałych (jak w klasycznej analizie korespondencji z dwiema zmiennymi) i stosowanie kryterium stopnia wyjaśniania bezwładności całkowitej przez inercje główne niewiele zmniejsza wymiar przestrzeni rzutowania w porównaniu z wymiarem przestrzeni rzeczywistych powiązań. Należy w takim przypadku sprawdzić, jaki wymiar rzutowania wskażą inne metody.

Jedną z metod określenia wymiaru przestrzeni rzutowania, opartego na analizie macierzy Burta lub znaczników, jest kryterium liczby cech zaproponowane przez Greenacre'a. Za optymalny wymiar K^* przestrzeni uznaje się największą liczbę k , dla której wartości własne spełniają warunek:

$$\lambda_k > \frac{1}{Q}, \text{ czyli } K^* = \max_{k \in \{1, \dots, K\}; \lambda_k > \frac{1}{Q}}.$$

Greenacre podaje także metodę „ulepszenia” współrzędnych otrzymanych w wyniku przekształcenia macierzy Burta w następujący sposób:

$$\tilde{F} = F^* \Gamma^{-1} \tilde{A},$$

gdzie: \tilde{A} – macierz diagonalna pierwszych K^* wartości własnych zmodyfikowanych według wzoru: $\tilde{\lambda}_k = \left(\frac{Q}{Q-1}\right)^2 \left(\sqrt{\lambda_k} - \frac{1}{Q}\right)^2$,

F^* – macierz pierwszych K^* współrzędnych kategorii z macierzy F .

Efekty przeprowadzonej analizy korespondencji najczęściej prezentuje się graficznie i na tej podstawie ocenia położenie punktów wobec centrum rzutowania względem innych punktów określających kategorie tej samej zmiennej lub względem punktu opisującego kategorię innej cechy. Jeżeli punkt znajduje się blisko układu współrzędnych, oznacza to, że wyznaczony dla niego profil jest zbliżony do profilu przeciętnego. Punkty położone daleko od centrum rzutowania potwierdzają hipotezę o zależności cech. Jeżeli dwa punkty opisujące różne kategorie tej samej zmiennej leżą blisko siebie, to ich profile są podobne i możliwe jest połączenie liczebności dwóch kategorii w jedną bez istotnego wpływu na wyniki dalszych badań. Analizując położenie punktów opisujących kategorie dwóch różnych zmiennych, można stwierdzić, że istnieją powiązania między kategoriami, jeżeli punkty są położone blisko siebie. Jeżeli zaś punkty leżą po przeciwnych stronach centrum rzutowania, to oznacza, że kategorie nie występują wspólnie, a częstość występowania obu kategorii jednocześnie wynosi prawie zero.

Jeżeli najlepszą formą prezentacji współwystępowania cech jest przestrzeń o wymiarze większym niż trzy, należy wybrać inną metodę analizy wyników. Do tego celu, w przestrzeni zarówno o niskim, jak i wysokim wymiarze, można zastosować metody klasyfikacji. Jako obiekty należy określić kategorie wszystkich analizowanych cech, zmiennymi zaś są wartości współrzędnych rzutowania każdej kategorii. Metody klasyfikacji są również użyteczne, gdy liczba wszystkich wariantów cech jest duża oraz rozrzut punktów na wykresie nie pozwala jednoznacznie wyodrębnić klas. W pracy przeprowadzono klasyfikację wyników analizy korespondencji za pomocą jednej z bardziej popularnych aglomeracyjnych metod – metody Warda [Ostasiewicz 1998].

3. Materiał badawczy

Źródło danych w badaniu stanowiły ankiety wypełniane przez 465 studentów kilku szczecińskich uczelni. Ankieta zawierała 24 pytania dotyczące zjawiska kupowania za pośrednictwem Internetu na aukcjach lub w sklepach internetowych. Na podstawie ankiet wyodrębniono grupę 339 osób, które robiły zakupy przez Internet i przyjęto następujący zestaw zmiennych i odpowiadających im kategorii:

CzZ – Częstość kupowania przez Internet – często (*cz*), średnio (*sr*), rzadko (*rz*).

CzS – Częstość szukania produktów przez Internet – często (*cz*), średnio (*sr*), rzadko (*rz*).

Pow – Ważność wymienionych powodów: możliwość zakupu towaru niedostępnego w okolicy, większy wybór towarów, niższe ceny, możliwość łatwiejszego porównania cen u różnych sprzedawców, całodobowy dostęp do sklepów/aukcji, brak konieczności wychodzenia z domu, dostawa do domu, możliwość zwrotu zakupionego towaru bez podania przyczyny, moda na zakupy w sieci – ważne (*waz*), średnio ważne (*sr*), mało ważne (*m*).

Deklarując ważność, respondenci mogli nadawać poszczególnym powodom wagi od 1 do 5, przy czym możliwe było nadanie wag dowolnej liczbie czynników skłaniających do zakupów przez Internet. Do kategorii „ważne” powody zakupów zaliczono odpowiedzi tych respondentów, którzy ustalili wagi o wartości mediany z przedziału $(3,5;5>$, „średnio ważne”, gdy mediana miała wartość z przedziału $(2; 3,5>$, „mało ważne” zaś dla mediany z przedziału $<0; 2>$.

Gdz – Miejsce najczęstszych zakupów – serwisy aukcyjne (*a*), sklepy internetowe (*s*), równie często na aukcjach, jak i w sklepach (*r*).

KorzP – Korzystanie kiedykolwiek z porównywarek cen, np. nokaut.pl, skapiec.pl, ceno.pl – tak (*t*), nie (*n*).

CzKorz – Częstość korzystania z porównywarki cen – często (*cz*), średnio (*sr*), rzadko (*rz*).

PowZ – Powody wyboru sklepu internetowego lub serwisu aukcyjnego, w którym dokonałeś zakupu – sprawdzenie rzetelności (*spr*), niska cena (*tanio*), reklama (*rekl*) (sklep reklamowany w telewizji, Internecie, wygodna obsługa strony internetowej, zamieszczone zdjęcia sprzedawanego egzemplarza, opcja „licytacja” lub „kup teraz”), inne (*inne*).

Z_1 – Częstość kupowania do tej pory przez Internet: kosmetyków, leków, preparatów dietetycznych, sprzętu sportowego i odżywek, sprzętu rehabilitacyjnego, artykułów erotycznych – często (*cz*), rzadko (*rz*).

Z_2 – Częstość kupowania do tej pory przez Internet: wycieczek, biletów lotniczych, biletów kolejowych, na koncerty, na mecze – często (*cz*), rzadko (*rz*).

Z_3 – Częstość kupowania do tej pory przez Internet: ubrań, bielizny, obuwia, biżuterii – często (*cz*), rzadko (*rz*).

Z_4 – Częstość kupowania do tej pory przez Internet: sprzętu komputerowego, oprogramowania, gier, telefonów komórkowych i akcesoriów, sprzętu AGD/TV, samochodów, części i akcesoriów samochodowych – często (*cz*), rzadko (*rz*).

Z_5 – Częstość kupowania do tej pory przez Internet: książek, komiksów, prasy, zabawek, artykułów biurowych i papierniczych – często (*cz*), rzadko (*rz*).

Z_6 – Częstość kupowania do tej pory przez Internet: mebli, wyposażenia domu i ogrodu, antyków, dzieł sztuki, alkoholu – często (*cz*), rzadko (*rz*).

Dost – Dostępność przedmiotów, które chciałbyś kupić przez Internet – tak (*t*), nie (*n*).

Pł – Preferowana forma płatności za zakupy przez Internet – karta płatnicza lub kredytowa (*ka*); przelew/transfer (*prz*); płatność przez SMS (*sms*); przy odbiorze (*odb*).

Trans – Preferowana forma transportu zakupionych produktów – przesyłka pocztowa zwykła (*zwyk*), przesyłka pocztowa priorytetowa (*prio*), przesyłka kurierska (*kur*), odbiór osobisty (*osob*).

Płeć – Płeć – kobieta (*k*), mężczyzna (*m*).

Ucz – Uczelnia – Politechnika Szczecińska (PS), Pomorska Akademia Medyczna (PAM), Uniwersytet Szczeciński (US).

Zam – Aktualne miejsce zamieszkania – dom rodzinny (*r*), poza domem (*poza*) (akademik, stacja).

Pr – Praca – tak (*t*), nie (*n*).

Styp – Stypendium – tak (*t*), nie (*n*).

4. Wyniki badania

Przed zastosowaniem wielowymiarowej analizy korespondencji zbadano, czy między zmienną określającą częstość kupowania przez Internet a pozostałymi zmiennymi istnieje zależność. Okazało się, że zmienna określająca Częstość kupowania (*CzZ*) przez Internet wykazuje zależność z następującymi zmiennymi: Częstość kupowania do tej pory przez Internet: ubrań, bielizny, obuwia, biżuterii (*Z_3*), Częstość kupowania do tej pory przez Internet: sprzętu komputerowego, oprogramowania, gier, telefonów komórkowych i akcesoriów, sprzętu AGD/TV, samochodów, części i akcesoriów samochodowych (*Z_4*), Częstość kupowania do tej pory przez Internet: książek, komiksów, prasy, zabawek, artykułów biurowych i papierniczych (*Z_5*), Częstość szukania produktów przez Internet (*CzS*), Ważność wymienionych powodów (*Pow*), Powody wyboru sklepu internetowego lub serwisu aukcyjnego, w którym dokonano zakupu (*PowZ*), Preferowana forma transportu zakupionych produktów (*Trans*), Miejsce najczęstszych zakupów (*Gdz*), Częstość korzystania z porównywarki cen (*CzKorz*), Korzystanie kiedykolwiek z porównywarek cen (*KorzP*), Preferowana forma płatności za zakupy przez Internet (*Pł*) oraz Płeć (*Płeć*). Otrzymany zestaw zmiennych został przyjęty w dalszej analizie.

W pierwszym kroku na podstawie kategorii wybranych do badania zmiennych otrzymano macierz Burta o wymiarach 37×37 . Fragment tej macierzy zamieszczono w tab. 2.

Wymiar rzeczywistej przestrzeni współwystępowania odpowiedzi na pytania wynosi 24. W następnym kroku sprawdzono, jak wartości własne przestrzeni o niższym wymiarze wyjaśniają całkowitą bezwładność. Wyniki zamieszczono w tab. 3, w której zawarto wartości własne λ_k (kwadraty wartości osobliwych γ_k), udział inercji głównych w całkowitej inercji (procent bezwładności λ_k/λ) oraz udział wartości własnych z wymiaru K w inercji całkowitej (procent skumulowany τ_K).

Tabela 2. Fragment macierzy Burta dla analizowanych zmiennych

	CzZ:sr	CzZ:cz	CzZ:rz	CzS:sr	CzS:cz	CzS:rz	Pow:sr	Pow:waz	Pow:m	Gdz:s	Gdz:a	...
CzZ:sr	143	0	0	75	38	30	69	30	44	30	76	...
CzZ:cz	0	163	0	68	92	3	71	54	38	17	104	...
CzZ:rz	0	0	33	8	5	20	10	7	16	13	13	...
CzS:sr	75	68	8	151	0	0	71	40	40	27	85	...
CzS:cz	38	92	5	0	135	0	60	39	36	18	82	...
CzS:rz	30	3	20	0	0	53	19	12	22	15	26	...
Pow:sr	69	71	10	71	60	19	150	0	0	29	87	...
Pow:waz	30	54	7	40	39	12	0	91	0	13	46	...
Pow:m	44	38	16	40	36	22	0	0	98	18	60	...
Gdz:s	30	17	13	27	18	15	29	13	18	60	0	...
Gdz:a	76	104	13	85	82	26	87	46	60	0	193	...
...

Źródło: obliczenia własne.

Tabela 3. Wartości osobliwe i wartości własne oraz stopień wyjaśnienia całkowitej bezwładności

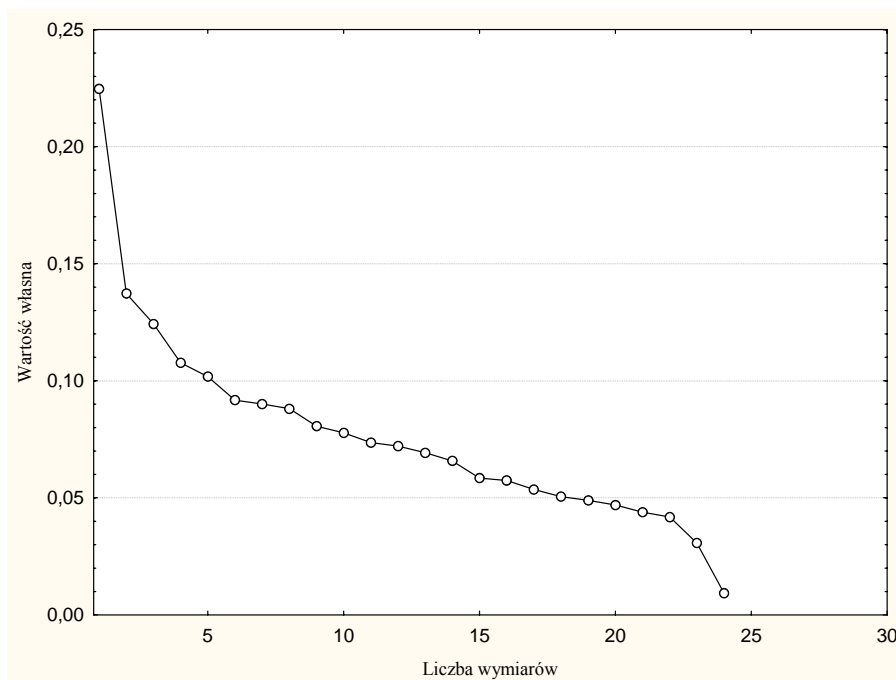
Liczba wymiarów K	Wartości osobliwe γ_k	Wartości własne λ_k	Procent bezwładności λ_k/λ	Procent skumulowany τ_K
1	0,4740	0,2247	12,1704	12,1704
2	0,3705	0,1373	7,4353	19,6057
3	0,3526	0,1243	6,7335	26,3392
4	0,3282	0,1077	5,8334	32,1726
5	0,3190	0,1018	5,5127	37,6853
6	0,3028	0,0917	4,9678	42,6531
7	0,3001	0,0901	4,8790	47,5321
8	0,2968	0,0881	4,7700	52,3021
9	0,2839	0,0806	4,3661	56,6682
10	0,2789	0,0778	4,2126	60,8808
11	0,2713	0,0736	3,9866	64,8674
12	0,2685	0,0721	3,9059	68,7733
13	0,2632	0,0693	3,7536	72,5270
14	0,2565	0,0658	3,5631	76,0901
15	0,2417	0,0584	3,1647	79,2547
16	0,2396	0,0574	3,1095	82,3642
17	0,2314	0,0536	2,9007	85,2649
18	0,2248	0,0505	2,7369	88,0018
19	0,2212	0,0489	2,6493	90,6511
20	0,2166	0,0469	2,5407	93,1919
21	0,2094	0,0438	2,3743	95,5662
22	0,2045	0,0418	2,2655	97,8317
23	0,1752	0,0307	1,6627	99,4943
24	0,0966	0,0093	0,5057	100,0000
		$\lambda = 1,8462$		

Źródło: obliczenia własne.

Według kryterium Greenacre'a jako najlepszy jest wybierany ten wymiar rzutowania kategorii zmiennych, w którym wartości własne spełniają warunek:

$$\lambda_{B,k} > \frac{1}{Q}. \text{ W analizowanym przypadku wartość ta wynosi } 0,0769, \text{ dla } Q=13. \text{ Na}$$

podstawie tab. 3 można stwierdzić, że są to bezwładności dla przestrzeni R^{10} , a łączna bezwładność w analizowanym przypadku wynosi 1,8462. Sporządzono także wykres wartości własnych (rys. 1) oraz dokonano modyfikacji wartości własnych według kryterium zaproponowanego przez Greenacre'a.



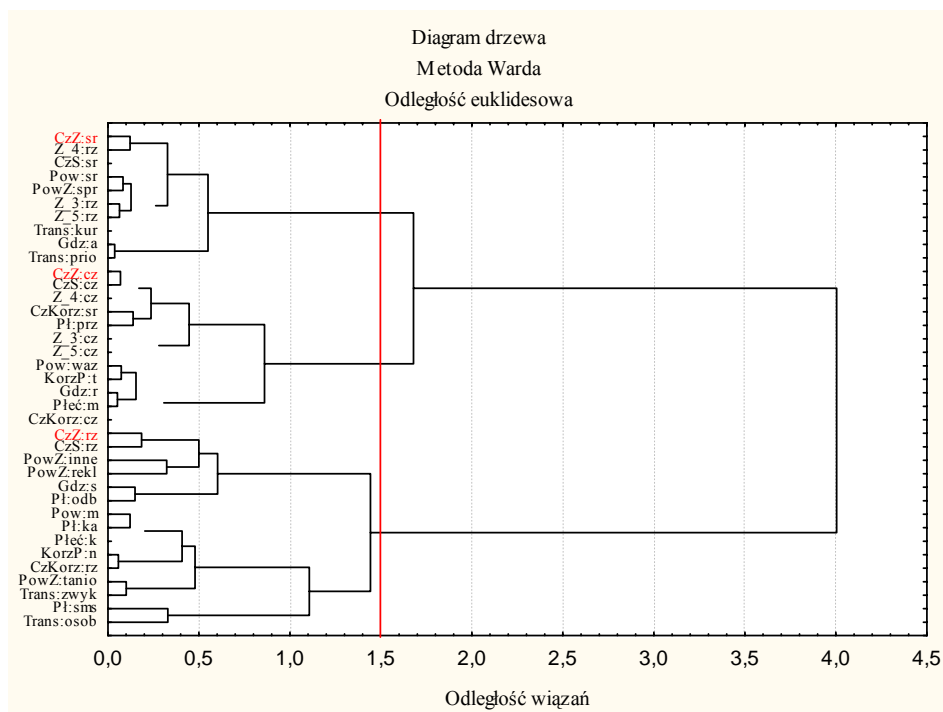
Rys. 1. Wykres wartości własnych

Źródło: opracowanie własne.

W tab. 4 zawarto wartości własne, wartości osobliwe, stopień wyjaśnienia całkowitej bezwładności po modyfikacji.

Po dokonaniu modyfikacji wartości własnych wymiar przestrzeni współwystępowania kategorii cech został ograniczony do R^4 . Do prezentacji graficznej przyjęto przestrzeń trójwymiarową. Trzy pierwsze wymiary odwzorowują 47,42% łącznej bezwładności. Zatem uwzględnienie pierwszych trzech wymiarów pozwala na wyjaśnienie prawie połowy bezwładności całkowitej. Na rys. 2 przedstawiono wyniki analizy korespondencji w przestrzeni trójwymiarowej z uwzględnieniem zmodyfikowanych wartości własnych.

Interpretacja tak zaprezentowanych wyników jest trudna, dlatego wykorzystano metodę Warda, za pomocą której wskazano powiązania między zmiennymi (rys. 3).



Rys. 3. Diagram podziału kategorii zmiennych za pomocą metody Warda

Źródło: opracowanie własne.

Rysunek 3 przedstawia dendrogram, na którym wyodrębniono trzy skupienia kategorii zmiennych. Pierwsza klasa obejmuje respondentów kupujących na aukcjach internetowych, którzy dokonywali zakupów przez Internet przeciętnie raz na pół roku, kupowali sprzęt komputerowy, oprogramowanie, gry, telefony komórkowe i akcesoria, sprzęt AGD/TV, samochody, części i akcesoria samochodowe (Z_4), ubrania, bieliznę, obuwie, biżuterię (Z_3) oraz książki, komiksy, prasę, zabawki, artykuły biurowe i papierowe (Z_5). Osoby z tej grupy kilka razy w miesiącu szukają produktów w Internecie. Jako powód wyboru aukcji podano rzetelność sprzedawcy. Preferowaną formą transportu zakupionych produktów był kurier lub przesyłka pocztowa priorytetowa. Druga klasa dotyczy mężczyzn, którzy często, czyli przeciętnie raz w miesiącu, kupują towary ze wszystkich grup Z_3, Z_4 oraz Z_5, zarówno w sklepach internetowych, jak i na aukcjach. Często, czyli co najmniej kilka razy w tygodniu, szukają towaru w Internecie, często korzystają z porównywarki cenowej, preferują płatność za pomocą przelewu. Trzecia grupa kategorii dotyczy kobiet, które kupują rzadko (najwyżej raz na dwa lata) w sklepach internetowych, kierując się przede wszystkim reklamą, rzadko

lub wcale nie korzystają z porównywarek cenowych, a powodem zakupu jest niska cena. Preferowaną formą płatności jest płatność za pomocą sms-a lub karty oraz przy odbiorze osobistym towaru.

5. Podsumowanie

W pracy przedstawiono wyniki badania powiązań między częstością e-zakupów a kategoriami zmiennych charakteryzujących badane zjawisko, jak np. miejsce dokonywania zakupów, powody dokonywania zakupów, najczęściej kupowane produkty, forma dokonywania płatności oraz transportu. Spośród zmiennych demograficzno-społecznych w badaniu uwzględniono jedynie płeć kupujących, pozostałe zmienne wykazywały niezależność ze zmienną określającą częstość zakupów przez Internet. Z badania wynika, że zakupy za pośrednictwem Internetu, zarówno w sklepach, jak i na aukcjach internetowych, dokonywane są najczęściej przez mężczyzn, którzy często kupują artykuły z różnych grup towarowych (np. ubrania, sprzęt AGD, RTV czy książki). Za zakupione towary płać najczęściej przelewem, przed zakupem korzystają z porównywarek cenowych, nie przywiązują wagi do formy transportu zakupionych artykułów. Natomiast rzadko kupują przez Internet kobiety, nie wskazując na rodzaj kupowanego towaru, a powodem zakupu jest niska cena. Wybierają różne formy transportu i płatności.

Literatura

- Gębka P., *Analiza statystyczna rynku i aukcji internetowych*, Praca magisterska, PS, Szczecin, 2008.
- Ostasiewicz W. (red.), *Statystyczne metody analizy danych*, Wydawnictwo AE, Wrocław 1998.
- Stanimir A., *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo AE, Wrocław 2005.
- Stanisz A., *Przystępny kurs z zastosowaniem Statistica PL na przykładach z medycyny*, T. 3. *Analizy wielowymiarowe*, Statsoft, Kraków 2007.

A CORRESPONDENCE ANALYSIS STUDY OF STUDENTS' ATTITUDES DEFINING ONLINE SHOPPING

Summary: In the work an attempt was made to analyse the frequency of buying by the Internet among students of the chosen universities in Szczecin. The aim of the study was an attempt to define the determinants affecting the frequency of doing shopping via the Internet. To reach the goal correspondence analysis was used, which allowed to define the connection between variables categories accepted to describe the phenomenon of buying by the Internet, e.g. place of doing shopping, reasons of doing shopping, mostly bought products, forms of making payments and transport. From among the demographic and social variables only sex of buyers was taken into account. The other variables showed independence with the variable defining the frequency of shopping online.