

EUGENIA CZERNYSZEWICZ

## ZASTOSOWANIE ANALIZY GŁÓWNYCH SKŁADOWYCH DO OPISU KONSUMENCKIEJ STRUKTURY JAKOŚCI JABŁEK

### Streszczenie

W pracy przedstawiono preferencje konsumentów lubelskich w zakresie cech wpływających na jakość jabłek. Do opisu preferencji zastosowano metodę analizy głównych składowych. Stwierdzono, że preferencje dotyczą głównie dwóch grup cech obejmujących charakterystyki związane z wyglądem zewnętrznym oraz przechowywaniem owoców, a ponadto z cechami odmianowymi, takim jak: smak oraz jędrność i soczystość miąższu. Analiza pozwoliła również zidentyfikować różnice w preferencjach kobiet i mężczyzn. Stwierdzono, że wielkość jabłek i odmiana różnicują preferencje kobiet i mężczyzn w niewielkim stopniu.

**Słowa kluczowe:** jabłka, kryteria jakościowe, analiza głównych składowych

### Wprowadzenie

Jabłka są ważnym elementem diety Polaków. W ostatnich latach ich spożycie w Polsce i Europie ma jednak tendencję spadkową [1, 4]. Znaczenie jabłek w naszym kraju wiąże się z faktem, że dominują one również w produkcji. Cechą charakterystyczną jest silne rozdrobnienie produkcji jabłek, które wpływa istotnie na jakość owoców dostępnych na rynku. Na zachowania nabywcze konsumentów wpływają w dużym stopniu cechy sensoryczne owoców, jak: wielkość, kształt, barwa skórki, smak, twardość miąższu i inne [2, 6, 8, 9]. W ostatnich latach rosnąca świadomość ekologiczna konsumentów wniosła dodatkowe wymagania jakościowe związane z bezpieczeństwem zdrowotnym. Z uwagi na fakt, że jabłka wysokiej jakości powinny spełniać przede wszystkim wymagania i zaspokajać potrzeby konsumentów, ważna staje się znajomość tych oczekiwań. Wiedza ta pomoże ukierunkować wysiłki, aby lepiej dostosować podaż do wymagań konsumentów.

Celem pracy była charakterystyka konsumenckiej struktury w zakresie kryteriów jakościowych jabłek. Opisu struktury jakości jabłek dokonano przy zastosowaniu me-

tody analizy głównych składowych, która pozwala przedstawić strukturę określoną wieloma zmiennymi za pomocą niewielkiej liczby czynników. Poddano weryfikacji także możliwość wykorzystania tej metody do charakterystyki struktury jakości jabłek w grupach konsumentów zróżnicowanych ze względu na płeć.

### **Material i metody badań**

Analizę wykonano na podstawie wyników badań ankietowych przeprowadzonych w roku 2004 wśród mieszkańców Lublina. Próba konsumentów liczyła 269 osób i była próbą kwotową odzwierciedlającą strukturę wiekową mieszkańców Lublina w czterech grupach: do 20 lat, 21-40 lat, 41-60 lat, powyżej 60 lat. W analizie uwzględniono dodatkowo jako zmienną demograficzną płeć konsumentów. W pracy wykorzystano wyniki oceny respondentów dotyczące kryteriów decydujących o wysokiej jakości jabłek. Poszczególne kryteria jakości jabłek oceniano za pomocą 5-stopniowej skali porządkowej. Cechy jakości jabłek obejmowały: świeży wygląd, smak, soczystość miąższu, jędrność miąższu, odmianę, wielkość owoców, brak obcych smaków i zapachów, brak objawów widocznych chorób czy szkodników, brak uszkodzeń mechanicznych oraz brak szkodliwych dla zdrowia pozostałości środków ochrony roślin.

Do opisu konsumenckiej struktury jakości jabłek zastosowano metodę analizy głównych składowych. Metodę tę zastosowano m.in. w pracy Péneau i wsp. [9], w której analizowano znaczenia świeżości jabłek dla konsumentów oraz związku pomiędzy różnymi cechami jabłek, a także w pracach: Konopackiej i wsp. [8], Karlsen i wsp. [7], Jaeger i wsp. [5] oraz Daillant-Spinnler i wsp. [3]. Analiza ta pozwala na przekształcenie danego zbioru cech (zmiennych), które są wzajemnie skorelowane w nowy układ cech (tzw. głównych składowych) wzajemnie nieskorelowanych, który jest porównywalny z układem wyjściowym. Analiza głównych składowych prowadzi więc do uproszczenia struktury danych. Charakteryzuje się tym, że obejmuje wariancję całkowitą zmiennych, wyjaśnia maksimum zmienności w zbiorze danych, główna składowa jest funkcją zmiennych pierwotnych oraz główne składowe są zawsze niezależne. Pierwszym etapem w tworzeniu modelu było zbudowanie macierzy korelacji między zmiennymi (cechami) pierwotnymi i wstępna analiza uzyskanych danych. Z uwagi na to, że uwzględnione w analizie zmienne były jednomianowe nie przeprowadzono wcześniej zalecanej standaryzacji zmiennych. W literaturze tematu uznaje się, że jeśli współczynniki korelacji są niskie (niższe od 0,3) to należy zaniechać dalszej procedury postępowania według analizy głównych składowych [11]. Jeśli zaś w macierzy korelacji każda zmienna koreluje wysoko (powyżej 0,3) z którąkolwiek z pozostałych to można przejść do określenia sposobu wyodrębniania głównych składowych. W analizie, wariancja nowej zmiennej (wyjaśniającej pewną ilość zmienności zmiennych pierwotnych) nazywana jest jej wartością własną. Do określenia liczby głównych składowych zastosowano metodę wartości własnej większej od jedności. Technika ustala-

nia liczby składowych na podstawie wartości własnych większych niż 1 nazwano „kryterium Kaisera” od nazwiska osoby, która zaproponowała ten sposób postępowania [12]. W metodzie tej zakłada się, że każdy czynnik powinien wyjaśniać zmienność co najmniej jednej zmiennej pierwotnej. Na nazwanie nowych zmiennych pozwalają tzw. ładunki czynnikowe, które wyrażają stopień nasycenia głównej składowej danym czynnikiem. Im większy jest współczynnik korelacji zmiennej ze składową, tym bardziej istotna jest ta zmienna dla danej głównej składowej. W dalszej części analizy ładunki czynnikowe poddano technice rotacji ortogonalnej *varimax*. Technika ta dzięki minimalizacji liczby zmiennych potrzebnych do wyjaśnienia danej zmiennej upraszcza interpretację wyników. W celu nazwania nowych zmiennych (głównych składowych) wyodrębniono te o najwyższych ładunkach czynnikowych względem danych składowych. W prezentowanej analizie wiersze macierzy danych  $X$  ( $n \times m$ ) odpowiadały próbie badanych konsumentów, a kolumny zmiennym objaśniającym konsumencką strukturę jakości jabłek. Obliczenia statystyczne wykonano w systemie SAS wersja 9.1.

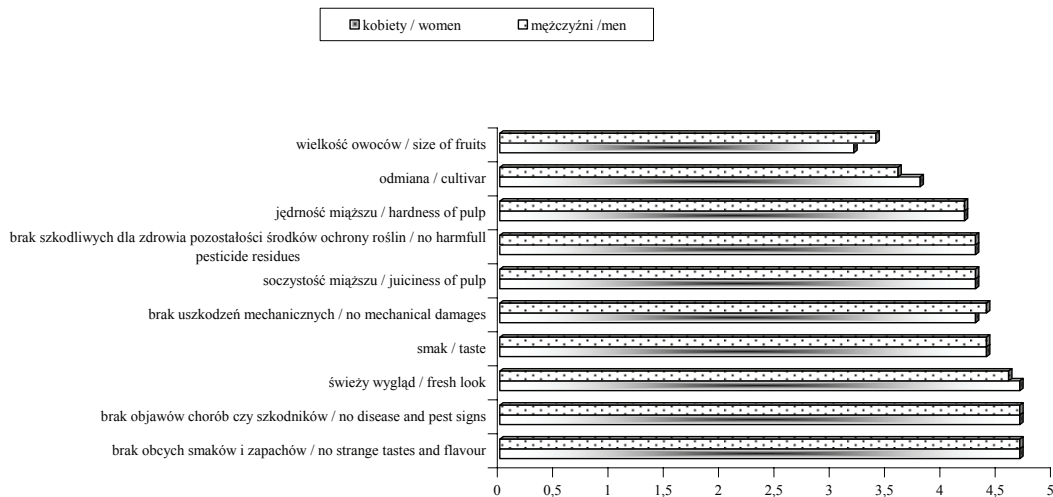
### Wyniki i dyskusja

Na wykresie (rys. 1) przedstawiono ważność cech jakościowych jabłek dla kobiet i mężczyzn. Cechy, takie jak wielkość owoców, brak uszkodzeń mechanicznych były wyżej oceniane przez mężczyzn niż przez kobiety, dla których ważniejszy był świeży wygląd i odmiana. Z danych Czernyszewicz [2] wynika, że współczynniki korelacji (Pearsona i Spearmana) pomiędzy płcią respondentów a wymienionymi cechami są niskie i statystycznie nieistotne. Analizując te dane trudno jest ustalić cechy, które jednoznacznie określają i odróżniają preferencje kobiet i mężczyzn. W tej sytuacji, aby lepiej scharakteryzować strukturę jakości jabłek określoną wieloma zmiennymi zastosowano metodę analizy głównych składowych.

W tab. 1. przedstawiono macierz współczynników korelacji między zmiennymi, którymi są parametry jakości jabłek. Pomiedzy wieloma parametrami korelacja jest większa od 0,3, co uzasadnia przeprowadzenie procedury polegającej na określeniu sposobu wyodrębniania głównych składowych i ich dalszej analizy.

W tab. 2. zestawiono wartości własne macierzy korelacji, będące wariancją głównych składowych oraz ich procentowy udział w zmienności zbioru danych.

Z przedstawionych danych wynika, że stosując kryterium wartości własnych większych od 1 do opisu preferencji całej próby wystarczą dwie pierwsze główne składowe, a dla grupy respondentów wyodrębnionych ze względu na płeć – trzy pierwsze główne składowe. Ze względu jednak na to, że składowe te wyjaśniają odpowiednio tylko około 52, 64 i 61% zmienności zawartej w danych uzasadnione byłoby zwiększenie liczby składowych do 4, aby co najmniej 70% zmienności było wyjaśnione przez główne składowe. Uwzględniając jednak aspekt interpretacji dalszą analizę przeprowadzono stosując opcję, która dopuszcza tylko ładunki większe od 1.



Rys. 1. Średnia ważność parametrów jakości jabłek dla kobiet i mężczyzn (w skali 1-5).

Fig. 1. Average importance of some quality parameters of apples for women and men (on a scale from 1 to 5).

W tab. 3. przedstawiono wartości tzw. ładunków czynnikowych, na podstawie których można określić, które zmienne (cechy) są najbardziej istotne dla danej głównej składowej. W interpretacji głównych składowych uwzględniono zmienne o najwyższych ładunkach czynnikowych. Zastosowanie rotacji ortogonalnej istotnie ułatwiło określenie zmiennych wyjaśniających główną składową. Dodatkowo na wykresie (rys. 2) przedstawiono położenie analizowanych zmiennych względem dwóch głównych składowych. Z obrazu wykresu wynika, że zmienna F oznaczająca cechę brak szkodliwych dla zdrowia pozostałości środków ochrony roślin w pierwszej składowej i zmienna A (wielkość owoców) w drugiej różnią się od pozostałych.

Zgodnie z przedstawionymi danymi konsumentką strukturę jakości jabłek w Lublinie można opisać za pomocą 2 grup cech, obejmujących większość badanych zmiennych. Pierwsza grupa, która wyjaśnia zmienność zestawu danych w ponad 33%, obejmuje charakterystyki związane z wyglądem zewnętrznym jabłek (świeży wygląd, brak uszkodzeń mechanicznych oraz objawów chorób czy szkodników) oraz przechowywaniem owoców, które wiążą się z brakiem obcych zapachów i smaków. Druga grupa cech określających jakość jabłek, wyjaśniająca ponad 18% zmienności danych, związana jest z cechami odmianowymi, takim jak smak oraz jędrność i soczystość miąższu. Pozostałe cechy, takie jak wielkość owoców i brak szkodliwych dla zdrowia pozostałości środków ochrony roślin objaśniają konsumentką strukturę jakości w znacznie mniejszym stopniu.

Tabela 1

Macierz korelacji parametrów jakości jabłek.

Matrix of correlation among the parameters of apple quality.

	A	B	C	D	E	F	G	H	I	J
Ogółem / Total										
A	1									
B	0,299	1								
C	0,233	0,389	1							
D	0,226	0,416	0,492	1						
E	0,218	0,313	0,503	0,504	1					
F	-0,001	0,015	0,083	0,142	0,172	1				
G	0,148	0,107	0,136	0,142	0,070	0,372	1			
H	0,172	0,121	0,197	0,345	0,193	0,317	0,493	1		
I	0,131	0,088	0,238	0,170	0,269	0,332	0,365	0,580	1	
J	0,206	0,012	0,140	0,164	0,710	0,239	0,371	0,488	0,524	1
Kobiety / Women										
A	1									
B	0,299	1								
C	0,233	0,389	1							
D	0,226	0,461	0,493	1						
E	0,218	0,313	0,503	0,504	1					
F	-0,001	0,015	0,084	0,142	0,172	1				
G	0,148	0,107	0,136	0,142	0,070	0,372	1			
H	0,172	0,121	0,197	0,345	0,193	0,317	0,493	1		
I	0,131	0,088	0,238	0,170	0,269	0,332	0,365	0,579	1	
J	0,206	0,012	0,140	0,164	0,170	0,239	0,371	0,488	0,524	1
Mężczyźni / Men										
A	1									
B	0,324	1								
C	0,241	0,508	1							
D	0,191	0,494	0,522	1						
E	0,265	0,446	0,559	0,586	1					
F	-0,013	0,039	0,038	0,069	0,101	1				
G	0,091	-0,024	0,028	0,134	0,046	0,387	1			
H	0,164	0,138	0,174	0,319	0,152	0,261	0,504	1		
I	0,035	0,164	0,305	0,203	0,231	0,332	0,371	0,531	1	
J	0,222	0,004	0,120	0,167	0,113	0,159	0,379	0,519	0,464	1

A – wielkość owoców / size of fruit; B – odmiana / cultivar; C – jędrność miąższu / firmness of parenchyma; D – smak / taste; E – soczystość miąższu / juiciness of parenchyma; F – brak szkodliwych dla zdrowia pozostałości środków ochrony roślin / no pesticide residues appearing harmful to health; G – brak uszkodzeń mechanicznych / no mechanical damages; H – świeży wygląd / fresh appearance; I – brak obcych smaków i zapachów / no strange tastes and flavours; J – brak objawów chorób czy szkodników / no symptoms of any disease and no signs of pests.

Tabela 2

Wariancje głównych składowych i ich procentowy udział w zmienności zestawu danych.  
Variances of principal components and their participation, expressed as a percent rate, in the variability of a data set.

PC	Ogółem / Total		Kobiety / Women		Mężczyźni / Men	
	Wartości własne macierzy korelacji Eigenvalues of correlation matrices	Udział % Participation expressed as a percent rate	Wartości własne macierzy korelacji Eigenvalues of correlation matrices	Udział % Participation expressed as a percent rate	Wartości własne macierzy korelacji Eigenvalues of correlation matrices	Udział % Participation expressed as a percent rate
1	3,341	33,4	3,305	33,1	3,404	34,0
2	1,846	51,9	2,056	53,6	1,680	50,8
3	0,963	61,5	1,007	63,7	1,043	61,3
4	0,826	69,8	0,831	72,0	0,928	70,6
5	0,695	76,7	0,653	78,5	0,677	77,3
6	0,565	82,4	0,578	84,3	0,654	83,9
7	0,546	87,8	0,484	89,1	0,593	89,8
8	0,480	92,6	0,406	93,2	0,485	94,6
9	0,444	97,1	0,381	97,0	0,323	97,9
10	0,293	100,0	0,294	100,0	0,214	100,0

Tabela 3

Wartości ładunków czynnikowych (głównych składowych) modelu po rotacji.  
Values of factor loads (of principal components) of the model after rotation.

Zmienne Variables	Ogółem / Total		Kobiety / Women			Mężczyźni / Men		
	Składowa 1	Składowa 2	Składowa 1	Składowa 2	Składowa 3	Składowa 1	Składowa 2	Składowa 3
	Component	Component	Component	Component	Component	Component	Component	Component
A	0,131	0,477	0,273	0,160	0,766	0,153	0,174	0,547
B	-0,052	0,724	0,762	-0,020	0,753	-0,171	0,300	0,749
C	0,107	0,759	0,806	0,089	0,053	0,089	0,698	0,289
D	0,154	0,778	0,780	0,176	0,036	0,061	0,727	0,385
E	0,144	0,722	0,808	0,082	0,024	0,288	0,805	-0,082
F	0,590	0,023	0,099	0,515	-0,543	0,618	0,228	-0,093
G	0,707	0,060	-0,039	0,758	-0,094	0,513	-0,110	0,646
H	0,783	0,210	0,155	0,800	0,114	0,740	0,131	0,296
I	0,765	0,170	0,257	0,713	-0,159	0,834	0,073	0,085
J	0,731	0,090	-0,005	0,743	0,360	0,769	0,090	0,061

A – wielkość owoców / size of fruit; B – odmiana / cultivar; C – jędrność miąższu / firmness of parenchyma; D – smak / taste; E – soczystość miąższu / juiciness of parenchyma; F – brak szkodliwych dla zdrowia pozostałości / no pesticide residues appearing harmful to health; G – brak uszkodzeń mechanicznych / no mechanical damages; H – świeży wygląd / fresh appearance; I – brak obcych smaków i zapachów / no strange tastes and flavours; J – brak objawów chorób czy szkodników / no symptoms of any disease and no signs of pests.



Uregulowania w zakresie m.in. wielkości owoców, braku obcych smaków i zapachów, widocznych objawów chorób czy szkodników oraz uszkodzeń mechanicznych owoców są zawarte w standardach jakości handlowej. Standardy jakości handlowej mające zastosowanie do jabłek zawiera rozporządzenie Komisji (WE) nr 85/2004 z 15 stycznia 2004 r. [10]. Jak wynika z badań cechy te, nie licząc wielkości owoców, mają także duże znaczenie w konsumenckiej ocenie jakości jabłek.

### **Wnioski**

1. Strukturę jakości jabłek konsumentów lubelskich można opisać za pomocą dwóch grup cech. Pierwsza grupa obejmuje charakterystyki związane z wyglądem zewnętrznym jabłek (świeży wygląd, brak uszkodzeń mechanicznych oraz objawów chorób czy szkodników) oraz właściwym przechowywaniem owoców, które oznacza m.in. brak obcych zapachów i smaków. Druga grupa związana jest z cechami odmianowymi, takim jak smak oraz jędrność i soczystość miąższu. Pozostałe badane cechy, jak wielkość owoców i brak szkodliwych dla zdrowia pozostałości środków ochrony roślin wyjaśniają konsumencką strukturę jakości jabłek w znacznie mniejszym stopniu.
2. Stwierdzono, że struktura jakości jabłek jest różna dla kobiet i mężczyzn. O ile zmienność struktury jakości dla kobiet w największym stopniu wyjaśniają cechy zależne od odmiany, to dla mężczyzn cechy związane z wyglądem zewnętrznym jabłek (świeży wygląd, brak uszkodzeń mechanicznych oraz objawów chorób czy szkodników) oraz cechami wewnętrznymi, determinowanymi przez właściwe przechowywanie i technologię, w tym właściwe stosowanie środków ochrony roślin. Wielkość jabłek i odmiana różnicowały strukturę jakości jabłek dla kobiet i mężczyzn w niewielkim stopniu.
3. Na podstawie wyników przedstawionej analizy można wnioskować o przydatności metody analizy głównych składowych do określenia ważności cech jakości jabłek dla konsumentów oraz charakterystyki struktury jakości jabłek dla różnych grup demograficznych i społeczno-ekonomicznych konsumentów.
4. Ze względu na mały udział poszczególnych charakterystyk jakości w wyjaśnieniu zmienności struktury lepszym kryterium doboru liczby głównych składowych, w porównaniu z metodą „wartości własnej większej od jedności”, jest metoda „procentu wariancji tłumaczonej przez główne składowe”. Jednak w tym przypadku, z powodu zbyt dużej liczby głównych składowych, mogą występować problemy z interpretacją wyników.



### Literatura

- [1] Czernyszewicz E.: Changes in fruit and fruit products consumption in the social-economic group of households during 1990-2005. *EJPAU*, 2008, 10, 4, <http://www.ejpau.media.pl/volume10/issue4/art-37.html>.
- [2] Czernyszewicz E.: Konsumentenckie spojrzenie na jakość jabłek. *Annales UMCS*, 2007, **sectio EEE, XVII (2)**, 70-82.
- [3] Dailliant-Spinnler B., MacFie H.J.H., Beyts P.K. & Hedderley D.: Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food Quality and Preference*, 1996, **2 (7)**, 113-126.
- [4] Ellinger W.: Konsumtrends bei Äpfeln und Apfelsorten. *Mat. Interpoma*, Bolzano 2006, s. 17-31.
- [5] Jaeger S.R., Andani Z., Wakeling I.N. & MacFie H.J.H.: Consumer preferences for fresh and aged apples: a cross-cultural comparison. *Food Quality and Preference*, 1998, **5 (9)**, 355-366.
- [6] Jesionkowska K., Konopacka D., Plocharski W.: The quality of apples – Preferences among Consumers from Skierniewice, Poland. *J. Fruit Orn. Plant Res.* 2006, **14**, 173-182.
- [7] Karlsen A.M., Aaby K., Sivertsen H., Baardseth P., Ellekjaer M.R.: Instrumental and sensory analysis of fresh Norwegian and imported apples. *Food Quality and Preference*, 1999, **10**, 305-314.
- [8] Konopacka D., Jesionkowska K., Rutkowski K., Plocharski W., Tomala K.: A comparison of the quality of well known and scab resistant apples in export and consumer evaluation. *Vegetable Crops Research Bulletin*, 2006, **65**, 185-194.
- [9] Péneau S., Hoehn E., Roth H.-R., Escher F., Nuessli J.: Importance and consumer perception of freshness of apples. *Food Quality and Preference*, 2006, **17**, 9-19.
- [10] Rozporządzenie Komisji (WE) nr 85/2004 z 15 stycznia 2004 r. ustanawiające normy handlowe mające zastosowanie do jabłek. *Dz. U. WE L Nr 13 z dnia 20.01.2004*, 3-18.
- [11] Sokołowski A., Sagan A.: Przykłady stosowania analizy danych w marketingu i badaniu opinii publicznej. <http://www.statsoft.pl/czytelnia/marketing/adwmarketingu.html>
- [12] Składowe główne i analiza czynnikowa. <http://www.statsoft.pl/textbook/stfacan.html>

#### THE APPLICATION OF PRINCIPAL COMPONENT ANALYSIS TO CHARACTERIZE A CONSUMER STRUCTURE OF APPLE QUALITY

##### S u m m a r y

In the paper, the preferences were presented of consumers in a city of Lublin regarding properties impacting the quality of apples. A 'PCA' method (Principal Component Analysis) was used to describe the preferences studied. It was found that those preferences mainly referred to two groups of properties connected with the external appearance of the fruit and the method of storing them; moreover, they also covered the characteristics of cultivars, such as: taste, firmness, and juiciness of parenchyma. With this Analysis applied, it was also possible to identify differences in the preferences of women and men. It was found that the size of apples and their cultivar only insignificantly differentiated the preferences of women and men.

**Key words:** apples, quality criteria, Principal Component Analysis ☒