

Metody statystyczne w segmentacji rynku

The statistical methods used in market segmentation

Katarzyna Dębkowska

Politechnika Białostocka, Wydział Zarządzania, Katedra Informatyki Gospodarczej i Logistyki

Abstract

This article aims to present selected statistical methods that can be used in market research, particularly in market segmentation. Through the study (analysis) we understand the market science-based recognition of the market mechanism, structure, condition and development of its components in order to create the conditions for decision making. Market analysis is to create the conditions for decision-making within the marketing management. Is posed by the requirement that the analysis was not only a diagnosis of current market conditions and its past states, but also to represent a vision for the future. This forces the use of appropriate research methodology. No doubt the methods that significantly meet the above requirements are posed statistical methods. The paper will be presented including multivariate statistical methods that allow consumers to explore their needs, preferences, behavior in the market. The article presents the following statistical methods: cluster analysis, classification trees, correspondence analysis. The study was conducted among clients lingerie market.

Keywords: market research, market segmentation, cluster analysis, classification trees, correspondence analysis

Wstęp

Badanie (analiza) rynku, to oparte na naukowych podstawach, rozpoznanie mechanizmu rynku, jego struktury, stanu i rozwoju jego elementów, w celu stworzenia przesłanek do podejmowania decyzji. Często spotykamy się z określeniem „badania marketingowe”. Występuje wiele definicji badań marketingowych. Ph. Kotler określa badania marketingowe jako systematyczne projektowanie, zbieranie, analizowanie i prezentowanie danych oraz wyników badań związanych istotnie ze specyficzną sytuacją marketingową przedsiębiorstwa. Głównym zadaniem badań mar-

ketingowych jest dostarczenie informacji decydom i innym użytkownikom pracującym w przedsiębiorstwie. Konieczność prowadzenia badań, wynika z potrzeb informacyjnych decydentów. Istotą badań marketingowych jest pomoc w podejmowaniu trafnych decyzji. Pojęcia badania rynkowe i badania marketingowe nie są więc synonimami, tylko określeniami bliskoznacznymi.

Tabela 1. Metody statystyczne i ich zastosowanie w badaniach rynku

Metoda statystyczna	Zastosowanie
Analiza regresji	<ul style="list-style-type: none"> • Określenie i prognozowanie udziału w rynku, wielkość sprzedaży • Rozpoznawanie związków konkurencyjnych (analiza cen, wydatków na reklamę i promocję dóbr konkurencyjnych)
Metody klasyfikacji	<ul style="list-style-type: none"> • Segmentacja rynku • Określenie pozycji produktu na rynku (pozycjonowanie i repozycjonowanie produktu) • Identyfikacja rynków testowych • Określania struktury rynku (wydzielanie względnie jednorodnych klas produktów)
Conjoint analysis	<ul style="list-style-type: none"> • Segmentacja rynku • Określenie cech produktów, które są dla konsumentów najistotniejsze • Badanie preferencji (nastawienia) konsumentów wobec produktów
Skalowanie wielowymiarowe	<ul style="list-style-type: none"> • Segmentacja rynku • Określanie pozycji produktu na rynku (wydzielanie względnie jednorodnych klas produktów) • Rozpoznawanie „luk” na rynku • Określania struktury rynku (wydzielanie względnie jednorodnych klas produktów) • Określenie cech produktów, które są dla konsumentów najistotniejsze
Drzewa klasyfikacyjne	<ul style="list-style-type: none"> • Segmentacja rynku • Określenie cech produktów, które są dla konsumentów najistotniejsze
Analiza czynnikowa	<ul style="list-style-type: none"> • Określenie pozycji produktu na rynku (pozycjonowanie i repozycjonowanie produktu) • Rozpoznawanie „luk” na rynku • Określenie cech produktu, które są dla konsumentów najistotniejsze
Analiza korespondencji	<ul style="list-style-type: none"> • Segmentacja rynku • Określenie pozycji produktu na rynku (pozycjonowanie i repozycjonowanie produktu)
Metody porządkowanie liniowego	<ul style="list-style-type: none"> • Określenie pozycji produktu na rynku (pozycjonowanie i repozycjonowanie produktu) • Badanie preferencji (nastawienia) konsumentów wobec produktów

Źródło: opracowanie własne na podstawie: Gatnar, Walesiak 2004.

Analiza rynku ma stworzyć przesłanki do podejmowania decyzji w ramach zarządzania marketingowego. Stawiane jest przy tym wymaganie, by analiza nie tylko była diagnozą stanów obecnych rynku i jego stanów przeszłych, ale także by stanowiła wizję przyszłości. Wymusza to stosowanie odpowiedniej metodyki badań. W badaniach rynku wykorzystuje się wiedzę z zakresu mikroekonomii, statystyki, ekonometrii, a także doświadczenie praktyczne.

Podstawowe zastosowania wybranych metod statystycznej analizy wielowymiarowej w badaniach rynku zaprezentowano w tabeli 1. Warto zwrócić uwagę na fakt, że często pojawiającym się zastosowaniem jest segmentacja rynku. Dalsza część artykułu zostanie poświęcona temu zagadnieniu, które jest niewątpliwie bardzo istotne, z punktu widzenia badań rynkowych.

Segmentacja polega na podziale niejednorodnej grupy obiektów (klientów) na grupy. Wszystkie osoby znajdujące się w tej samej grupie uważane są za podobne do siebie, osoby znajdujące się w różnych grupach są różne. Dzięki tego typu podziałowi, nie musimy już określać tylu strategii, ilu mamy klientów. Wystarczy jeśli dany sposób postępowania przypiszemy do całej grupy (segmentu) podobnych osób.

Liczba segmentów zależy od tego, jak zróżnicowani są nasi klienci, w praktyce jednak nie spotyka się firm, które z powodzeniem zarządzają i komunikują się z 10 lub więcej segmentami (Berson, 2000). Oczywiście segmentacja ma sens jedynie wtedy, gdy planujemy wyszczególnione grupy traktować w odmienny sposób.

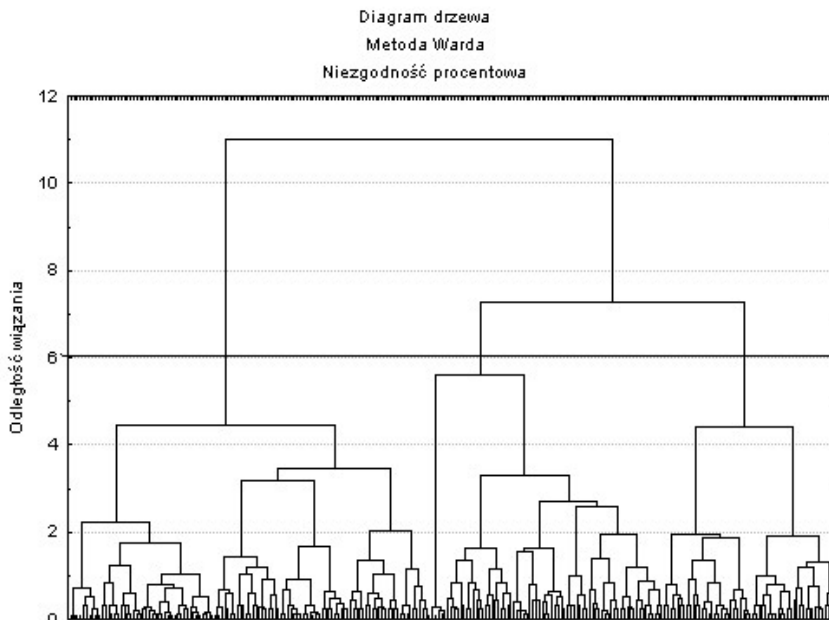
Wykorzystanie wybranych metod statystyki wielowymiarowej w segmentacji rynku na przykładzie badań klientek rynku bielizny

Jako statystyczne metody segmentacji rynku zostaną zaprezentowane metody analizy skupień, drzewa klasyfikacyjne oraz analiza korespondencji. Metody te zostaną każdorazowo wykorzystane do wyłonienia segmentów klientek na rynku bielizny. W tym celu zostały wykorzystane wyniki ogólnopolskiego badania przeprowadzonego na próbie ponad 1000 kobiet. Badanie odbywało się w miesiącach kwiecień – maj 2010 r. i zostało przeprowadzone na potrzeby Podlaskiego Klastra Bielizny.

Aby zaprezentować możliwość wykorzystania metod analizy skupień w segmentacji rynku, dokonano podziału kobiet ze względu na czynniki decydujące o zakupie bielizny. Użyto w tym celu dwóch metod analizy skupień: hierarchicznej metody Warda oraz niehierarchicznej metody k -średnich (z zadaną liczbą skupień wynikającą z dendrogramu metody Warda). Jako cechy diagnostyczne w procesie

tej segmentacji brano pod uwagę: funkcjonalność, modę, jakość, cenę, komfort, materiał, markę, krój oraz wygląd. Klientki oceniły oddziaływanie poszczególnych cech na zakup bielizny w skali od 1 do 5, gdzie: 1 – oznacza brak znaczenia danego czynnika, 2 – słabe znaczenie, 3 – umiarkowane znaczenie, 4 – duże znaczenie, 5 – bardzo duże znaczenie

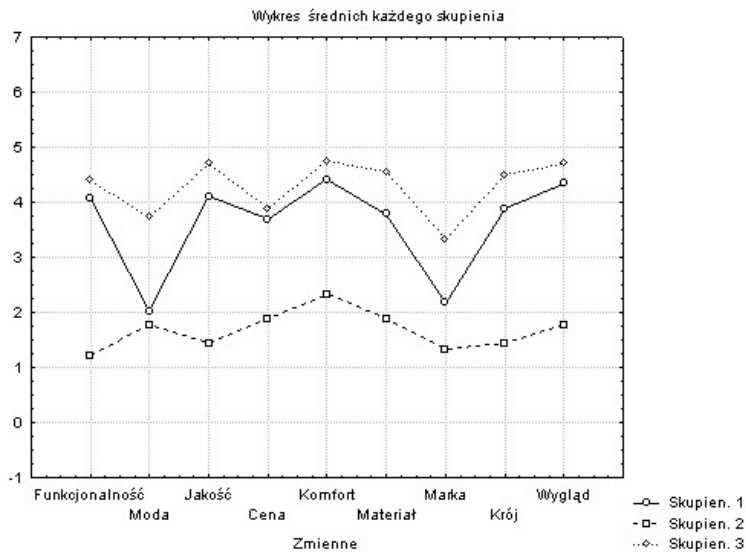
Używając hierarchicznej metody Warda otrzymano dendrogram, na podstawie którego wybrano trzy skupienia (rys.1).



Źródło: opracowanie własne przy użyciu pakietu *STATISTICA PL*.

Rys. 1. Podział kobiet ze względu na czynniki decydujące o zakupie bielizny za pomocą metody Warda

Po wyłonieniu grup kobiet o podobnych ocenach poszczególnych czynników wyznaczono średnie poziomy tych ocen w poszczególnych skupieniach (rys. 2.).



Źródło: opracowanie własne przy użyciu pakietu *STATISTICA PL*.

Rys. 2. Średnie poziomy ocen czynników decydujących o zakupie bielizny w wyróżnionych skupieniach

W skupieniu 1 znalazły się kobiety, dla których największe znaczenie miała użyteczność, niskie wartości zaś przyjmuje moda i marka, czyli psychologiczne czynniki decyzyjne zakupu. Można zatem do tego segmentu zaliczyć kobiety o dużym znaczeniu użyteczności. W skupieniu 2 znalazły się konsumentki mało wymagające, dla których wartości poszczególnych czynników osiągały najniższe poziomy. Zatem do tego segmentu zaliczyć można kobiety mało wymagające. Natomiast skupienie 3 stanowiły konsumentki o najwyższych poziomach badanych cech diagnostycznych, można je zatem uznać za klientki najbardziej wymagające.

Jako kolejny sposób segmentacji można wykorzystać metodę drzew klasyfikacyjnych. Jak wskazuje już sama nazwa drzewo klasyfikacyjne jest graficzną prezentacją modelu:

$$y = \sum_{k=1}^K a_k I(x_i \in R_{k_i}),$$

który powstaje w wyniku podziału wielowymiarowej przestrzeni zmiennych X^m , gdzie znajdują się klasyfikowane obiekty $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$. Wspomnianego podziału dokonuje się po to, by uzyskać rozłączne fragmenty (segmenty) tej przestrzeni R_k , w których znajdują się obiekty należące do tej samej klasy, którą reprezentuje

zmienna zależna y , przyjmująca wartości $l = 1, \dots, L$ (Gatnar, Walesiak, 2004). W modelu \mathbf{x}_i oznacza wielowymiarową obserwację, α_k - parametry modelu, I zaś zwykłą funkcję wskaźnikową:

$$I(q) = \begin{cases} 1 & \text{gdy } q \text{ jest prawdziwe} \\ 0 & \text{w przeciwnym wypadku} \end{cases}$$

Problem oddzielania obiektów należących do różnych klas jest rozważany na gruncie statystycznej analizy wielowymiarowej w ramach klasyfikacji.

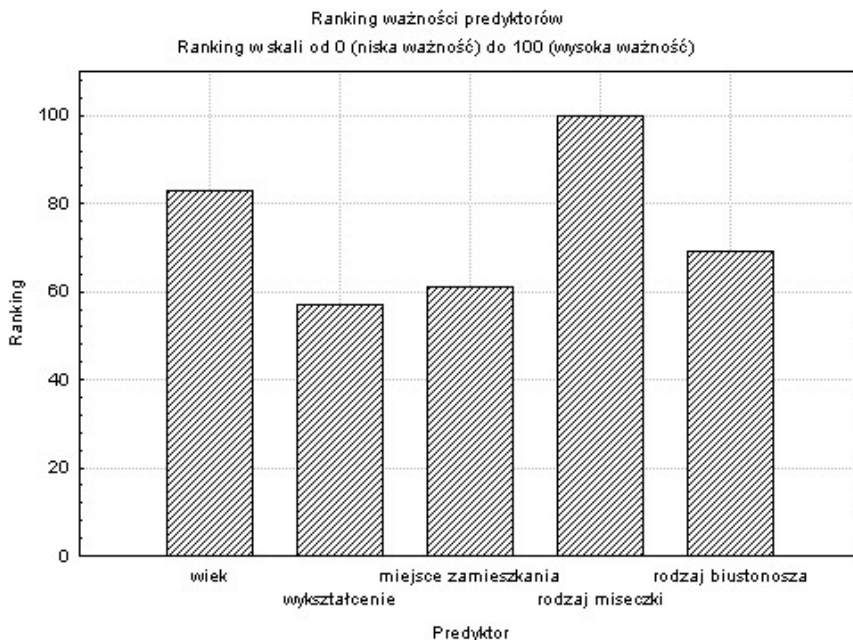
Drzewa klasyfikacyjne wykorzystano przy segmentacji klientek rynku bielizny. Jako zmienną objaśnianą przyjęto kwotę wydawaną na biustonosz, która rozpatrywana była w pięciu przedziałach: do 30 zł, 30-50 zł, 50-75 zł, 75-100 zł, powyżej 100 zł. Za zmienne objaśniające w tej segmentacji uznano: wiek, wykształcenie, miejsce zamieszkania, rozmiar miseczki biustonosza, rodzaj najczęściej kupowanego biustonosza (tabela 2)

Tabela 2. Zmienne wykorzystywane podczas segmentacji klientek rynku bielizny za pomocą drzew klasyfikacyjnych

Zmienna objaśniająca	Kategorie zmiennych	Odsetek badanej zbiorowości
Wiek	do 25 lat	21,3
	26-35 lat	28,9
	36-45 lat	34,6
	powyżej 45 lat	15,2
Wykształcenie	poniżej średniego	7,1
	średnie	42,9
	wyższe	50,0
Miejsce zamieszkania	do 15 tys. mieszkańców	5,1
	15-50 tys. mieszkańców	8,0
	50-300 tys. mieszkańców	43,1
	powyżej 300 tys. mieszkańców	43,8
Rozmiar miseczki biustonosza	A	5,7
	B	31,9
	C	34,1
	D	17,6
	E	4,4
	F	3,1
	G	2,6
Rodzaj najczęściej kupowanego biustonosza	pushup	40,7
	półsztywniany	41,7
	miękki	17,6

Źródło: opracowanie własne.

W trakcie analizy za pomocą drzew klasyfikacyjnych możemy dokonać rankingu zmiennych objaśniających pod względem ich ważności (rys. 3). Pierwsze miejsce w tym rankingu zajął rodzaj miseczki, dość ważnym w klasyfikacji jest też wiek klientek. Najniższe miejsca w rankingu zajęło wykształcenie oraz miejsce zamieszkania klientki.



Źródło: opracowanie własne przy użyciu pakietu *STATISTICA PL*.

Rys. 3. Ranking ważność zmiennych objaśnianych w segmentacji klientek rynku bielizny za pomocą drzew klasyfikacyjnych

Zastosowanie drzew klasyfikacyjnym pozwoliło na wyłonienie następujących głównych segmentów klientek (por. rys. 4):

Segment 1 (n=103) to klientki z wykształceniem średnim, w wieku poniżej 25 lat oraz powyżej 35 lat, noszące rozmiar miseczki A i C wydające na biustonosz od 30 do 50 zł.

Segment 2 (n=96) to klientki z wykształceniem wyższym, noszące rozmiar miseczki B, C, D, E, w wieku powyżej 25 lat, zamieszkałe w miejscowościach powyżej 50 tys. mieszkańców, kupujące biustonosze półuszywniane wydające na biustonosz od 50 do 75 zł.

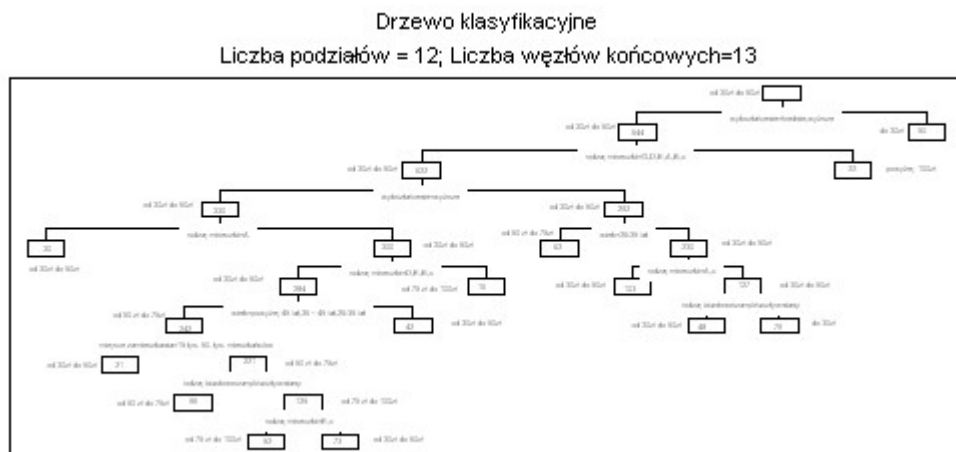
Segment 3 (n=79) to klientki z wykształceniem średnim, w wieku poniżej 25 lat oraz powyżej 35 lat, noszące rozmiar miseczki B, D, E, F, G preferujące biustonosz typu pushup i miękkie wydające na biustonosz do 30 zł.

Segment 4 (n=73) to klientki z wykształceniem wyższym, noszące rozmiar miseczki C, E, w wieku powyżej 25 lat, zamieszkałe w miejscowościach powyżej 50 tys. mieszkańców, kupujące biustonosze miękkie i pushup wydające na biustonosz od 30 do 50 zł.

Segment 5 (n=62) to klientki z wykształceniem średnim w wieku 26-35 lat, noszące rozmiar miseczki A, B, C, D, E, G wydające na biustonosz od 50 do 75 zł

Segment 6 (n=52) to klientki z wykształceniem wyższym, noszące rozmiar miseczki C, E, w wieku powyżej 25 lat, zamieszkałe w miejscowościach powyżej 50 tys. mieszkańców, kupujące biustonosze miękkie i pushup wydające na biustonosz od 75 do 100 zł.

Segment 7 (n=50) to klientki z wykształceniem poniżej średniego wydające na zakup biustonosza do 30 zł.



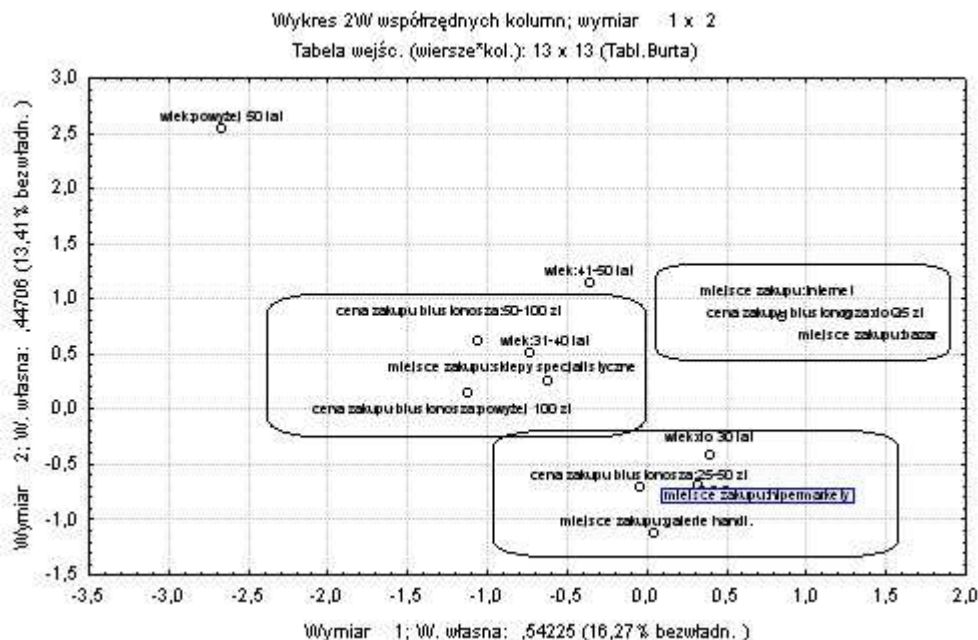
Źródło: opracowanie własne przy użyciu pakietu *STATISTICA PL*.

Rys. 4. Podział klientek rynku bielizny na segmenty za pomocą drzewa klasyfikacyjnego

Kolejną metodą statystyczną, która może mieć zastosowanie w segmentacji rynku jest analiza korespondencji, która jako eksploracyjna technika analizy tablic kontyngencji, zmierza do odtworzenia odległości, pomiędzy punktami reprezentującymi wiersze i/lub kolumny w przestrzeni o mniejszej liczbie wymiarów, przy jednoczesnym zachowaniu, jak największej ilości pierwotnej informacji. Głównym celem jest graficzna prezentacja zależności, między zmiennymi nominalnymi

w przestrzeni jedno-, dwu- lub trójwymiarowej. Analiza wykresu powinna w efekcie prowadzić do nadania interpretacji merytorycznej uzyskanym wymiarom, a w dalszej konsekwencji – zezwolić na wnioskowanie o związkach zachodzących między zmiennymi (Stanisz, 2007). Zastosowanie analizy korespondencji pomaga wykryć zależności, między zmiennymi jakościowymi, a szczególnie związki w interesującej nas zależności pomiędzy przynależnością do skupienia od reszty czynników (Panek, 2009).

Jako zmienne diagnostyczne w procesie segmentacji klientek rynku bielizny za pomocą analizy korespondencji, brano pod uwagę: miejsce zakupu bielizny (sklepy specjalistyczne z bielizną (poza centrami handlowymi), galerie handlowe, hipermarkety, supermarkety, Internet, bazary), przedział cenowy kupowanego biustonosza (do 25 zł, 25 zł- 50 zł, 50zł – 100 zł, powyżej 100 zł) oraz wiek (do 30 lat, 31-40 lat, 41-50 lat, powyżej 50 lat). Na podstawie przeprowadzonej analizy wyodrębniono następujące segmenty (rys. 5)



Źródło: opracowanie własne przy użyciu pakietu STATISTICA PL.

Rys. 5. Segmentacja klientek rynku bielizny za pomocą analizy korespondencji

Segment 1 to klientki w wieku 31-40 lat, wydające na biustonosz powyżej 50 zł, częściej wybierające sklepy specjalistyczne jako miejsce zakupu bielizny. Segment 2 to klientki wydające na biustonosz do 25 zł, częściej wybierające Internet albo bazar jako miejsce zakupu bielizny. Segment 3 to klientki w wieku do 30 lat, wydające na biustonosz od 25 do 50 zł, częściej wybierające hipermarkety lub galerie handlowe jako miejsce zakupu bielizny.

Wnioski

Pokazana segmentacja klientek jest alternatywą dla tradycyjnej segmentacji ze względu na wiek, czy dochód klientów. Biorąc pod uwagę złożoność czynników wpływających na proces decyzyjny konsumentek rynku bielizny, nie wystarczająca jest segmentacja w oparciu tylko o cechy podstawowe. Należy brać pod uwagę wiele cech wpływających na zakup bielizny, a pokazana segmentacja umożliwia badanie tych cech jednocześnie.

Przeprowadzone podziały klientek przy wykorzystaniu wybranych metod statystyki wielowymiarowej powinny stanowić początek realizacji segmentacji rynku. Decyzję o przekształceniu wyodrębnionych klas w segmenty podejmuje się na podstawie spełnionych przez nie kryteriów efektywności segmentacji rynku. Należy zatem sprawdzić wymierność segmentu, czy istnieje możliwość pomiaru wielkości siły nabywczej i profilu segmentu. Nie bez znaczenia jest wielkość segmentu, dostępność i stabilność segmentu w czasie. Trzeba też określić, czy możliwe jest sformułowanie skutecznego programu marketingowego dla obsługi segmentu. I wreszcie, należy wziąć pod uwagę rozróżnialność segmentów, które w odmienny sposób reagują na poszczególne elementy marketingu-mix.

Zaproponowane metody statystyczne w segmentacji rynku, powinny zatem stanowić punkt wyjścia do bardziej szczegółowych analiz. Należy zauważyć, że w dobie rozwoju technologii i możliwości wykorzystywania statystycznych programów komputerowych, analiza wielowymiarowa może bardzo użyteczne narzędzie w badaniach rynku.

Piśmiennictwo

1. Berson A., Smith S., Thearling K., 2000. *Building Data Mining Applications for CRM*. McGraw Hill, New York.
2. Hill N., Alexander J., 2003. *Pomiar satysfakcji i lojalności klientów*. Oficyna Wydawnicza, Kraków.
3. Gatnar, M. Walesiak E., (red.) 2004. *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*. AE, Wrocław.
4. Panek T., 2009. *Statystyczne metody wielowymiarowej analizy porównawczej*, Szkoła Główna Handlowa, Warszawa 2009.
5. Stanisław A., 2007. *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny, Tom 3. Analizy wielowymiarowe*. StatSoft, Kraków.