

Iwona Forys

Uniwersytet Szczeciński

WIELOWYMIAROWA ANALIZA CECH MIESZKAŃ SPRZEDAWANYCH NA RYNKU WARSZAWSKIM W BADANIU CZASU TRWANIA OFERTY W SYSTEMIE MLS

Streszczenie: Przedmiotem analizy jest warszawski rynek mieszkaniowy. Negatywnie zwerfikowano hipotezę o stałości cech wpływających na płynność, niezależnie od sytuacji na rynku mieszkaniowym. W tym celu wykorzystano funkcję dyskryminacji. Wykazano, że funkcje dyskryminacyjne w sposób zadowalający klasyfikują transakcje do trzech zaproponowanych grup płynności: niska, wysoka i średnia.

Słowa kluczowe: płynność, rynek mieszkaniowy, klasyfikacja.

1. Wstęp

Powszechnie uznaje się, że rynek nieruchomości jest rynkiem lokalnym, co oznacza, że lokalizacja nieruchomości determinuje preferencje nabywców, okoliczności i parametry zawieranych umów kupna-sprzedaży [Kucharska-Stasiak 2006]. Większość badaczy koncentruje się jednak na modelowaniu ceny transakcyjnej, a równie interesującym zjawiskiem jest tempo sprzedaży mieszkań (płynność) ze względu na cykle koniunkturalne. Można przyjąć, że na rynku popytowym (faza koniunktury) tempo sprzedaży powinno być wysokie, natomiast na rynku podażowym (faza dekonunktury) niskie.

Płynność na rynku mieszkaniowym można zdefiniować jako czas pomiędzy zgłoszeniem oferty sprzedaży a momentem zawarcia ostatecznej umowy przeniesienia praw do nieruchomości (wskaźnik mierzony w dniach), poprzedzonej czasem niezbędnym do przygotowania tej transakcji, zawierającej cenę adekwatną do wartości rynkowej mieszkania. W artykule wprowadzono zmodyfikowaną definicję płynności aktywów na rynku mieszkaniowym, którą można utożsamiać z „czasem trwania oferty” w systemie oferowania mieszkań. Wskaźnikiem płynności jest liczba dni, jaka upływa od zgłoszenia oferty w bazie systemu do jej wycofania w

wyniku znalezienia kupca. Można również badać związek płynności mieszkań z ich cechami oraz preferencjami kupujących.

W artykule podjęta została próba zweryfikowania dwóch hipotez: (1) na tempo sprzedaży mieszkania w systemie MLS ma wpływ zestaw tych samych cech mieszkań niezależnie od fazy cyklu koniunkturalnego na rynku mieszkaniowym oraz (2) funkcja dyskryminacyjna pozwala w sposób jednoznaczny zakwalifikować mieszkania do grupy o określonej płynności. Celem artykułu jest wyodrębnienie parametrów związanych z cechami mieszkania, które decydują w różnym stopniu o długości trwania oferty sprzedaży na rynku wtórnym. Uzyskane wyniki mogą być pomocne w kształtowaniu polityki inwestycyjnej firm deweloperskich oraz w ocenie preferencji nabywców mieszkań.

2. Analiza dyskryminacyjna w badaniu rynkowych ofert mieszkaniowych

Analiza dyskryminacji stosowana jest do rozstrzygnięcia, które zmienne wyróżniają (dyskryminują) grupy wielowymiarowych obiektów o znanych parametrach przy możliwie minimalnych błędach klasyfikacji [Stanisz 2007]. Weryfikacja założeń funkcji dyskryminacyjnej oznacza sprawdzenie, czy każda zmienna ma rozkład normalny, a macierze wariancji (kowariancji) są homogeniczne w grupach oraz czy korelacja między średnimi i wariancjami jest wysoka. Należy również zwrócić uwagę na to, czy zmienne wykorzystane do analizy nie są całkowicie redundantne oraz czy najmniej liczna grupa ma kilka razy więcej przypadków niż liczba zmiennych dyskryminujących (najczęściej cztery lub pięć). Usunąć należy również zmienne odstające ze względu na wrażliwość na nie proponowanej metody.

Do analizy najczęściej wykorzystuje się liniową funkcję dyskryminacyjną D_{kj} postaci [Gatnar 2008; Stanisz 2007]:

$$D_{kj} = \alpha_o + \alpha_1 x_{1kj} + \dots + \alpha_p x_{pkj},$$

gdzie: D_{kj} – wartości kanonicznej funkcji dyskryminacyjnej dla k -tego przypadku ($k = 1, \dots, n$) w j -tej grupie ($j = 1, \dots, g$),

x_{ikj} – wartość i -tej zmiennej ($i = 1, \dots, p$) dla k -tego przypadku w j -tej grupie,

α_i – współczynniki kanonicznej funkcji dyskryminacyjnej

p – liczba zmiennych dyskryminacyjnych, n – liczebność próby, g – liczba grup.

Przesłanki merytoryczne decydują o wyborze zmiennej grupującej, która pozwala wstępnie zakwalifikować dany przypadek do jednej z grup. Natomiast wybór zestawu zmiennych diagnostycznych (x_{ikj}) determinują przesłanki merytoryczne oraz dostępne informacje. Zredukowany zestaw zmiennych, spełniających założenia funkcji dyskryminacyjnej, pozwala na estymację jej parametrów (współczynniki surowe) i wyznaczenie wartości funkcji. Z kolei standaryzacja surowych współ-

czynników funkcji ułatwia interpretację (ocenę) udziału każdej funkcji w różnicowaniu grup.

Wybór reguły klasyfikacyjnej umożliwia klasyfikację przypadków oraz weryfikację zgodności przyjętej klasyfikacji. Jedną z procedur klasyfikacyjnych jest zaproponowana przez Fishera [Stanisz 2007] liniowa kombinacja zmiennych dyskryminacyjnych, wyznaczana odrębnie dla każdej grupy w postaci funkcji liniowej:

$$K_i = \beta_{i0} + \beta_{i1}x_1 + \dots + \beta_{ip}x_p,$$

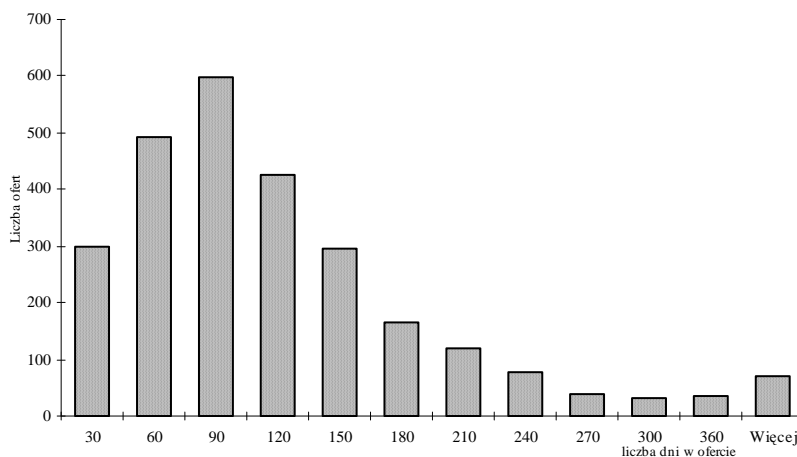
gdzie β_{ij} to współczynniki zmiennych dyskryminacyjnych funkcji klasyfikacyjnych dla kolejnych grup ($j = 0, \dots, p$). Każdy przypadek jest klasyfikowany do grupy, dla której K_i przyjmuje wartość największą. Etap klasyfikacji kończy weryfikacja zgodności klasyfikacji z prawdopodobieństwem *a priori* proporcjonalnym do wielkości grup, jednakowym oraz z innymi wagami, zaproponowanymi przez badacza. Podsumowaniem klasyfikacji mogą być macierze klasyfikacji, zawierające informacje na temat liczby i odsetka przypadków poprawnie sklasyfikowanych, oraz graficzna interpretacja rozrzutu wartości kanonicznych.

3. Wykorzystanie funkcji dyskryminacji w badaniu czasu trwania oferty mieszkaniowej w warszawskim systemie MLS

Do badania pozyskano zbiór 2646 ofert mieszkań w zabudowie wielorodzinnej z terenu miasta stołecznego Warszawa, które w latach 2003-2009 znalazły nabywców. Jest to badanie pełne, przeprowadzone na zasobie informacji zgromadzonych w systemie MLS WSPON¹. Pominięto oferty, które zostały wycofane przez zgłaszającego lub dotyczyły mieszkań niesprzedanych. Każdą z ofert opisuje zestaw kilkunastu cech jakościowych i ilościowych, w tym czas trwania oferty w systemie. Ze względu na postawioną hipotezę, w badaniu przyjęto czas trwania oferty w systemie MLS jako zmienną grupującą (X_1).

Ze względu na prawostronną asymetrię rozkładu zmiennej grupującej (rys. 1) można było przyjąć dolny i górny kwartył jako wartości wyznaczające granice poszczególnych grup, ostatecznie jednak przyjęto podział, który wynikał z opinii powszechnej wśród ekspertów na rynku nieruchomości: płynność wysoka (do 60 dni), płynność średnia (od 61 do 90 dni) oraz płynność niska (ponad 91 dni). Podział na trzy grupy zwiększył moc dyskryminacyjną modelu w stosunku do podziału na dwie grupy (spadła wartość statystyki Lambda Wilksa).

¹ System współpracy i wymiany informacji pomiędzy pośrednikami w obrocie nieruchomościami, obejmującego ponad 500 biur obrotu nieruchomościami, przede wszystkim z regionu mazowieckiego, prowadzony przez Warszawskie Stowarzyszenie Pośredników w Obrocie Nieruchomościami (WSPON).



Rys. 1. Rozkład wskaźnika płynności mieszkań na rynku warszawskim w latach 2003-2009 (w dniach) według MLS

Źródło: opracowanie własne na podstawie danych z systemu MLS.

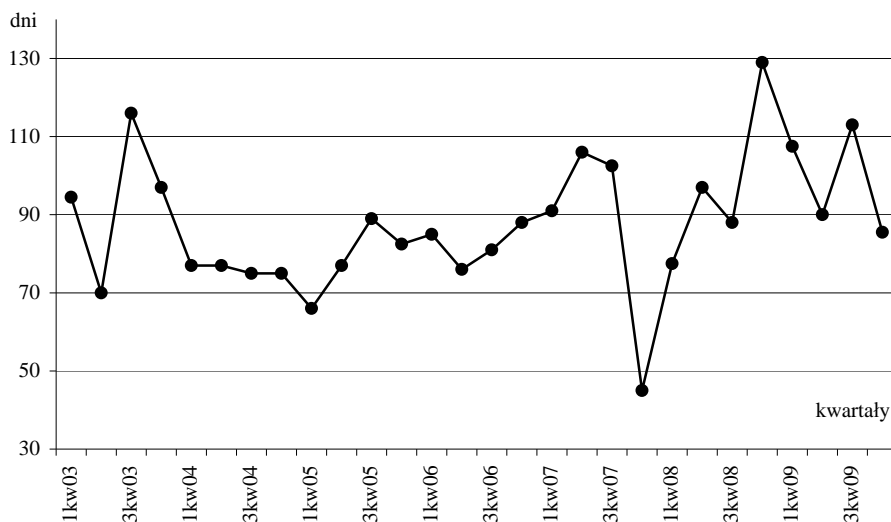
Do wyboru zmiennych diagnostycznych posłużył wyjściowy zbiór 99 cech opisujących każdą ofertę, który zweryfikowano ze względu na kompletność danych, pozostawiając do dalszej analizy 32 zmienne dotyczące lokalizacji i sąsiedztwa mieszkania, budynku, w którym znajduje się lokal, oraz charakterystyki udogodnień w lokalu, w tym zmienne z zakresu bezpieczeństwa. Z dostępnych danych wybrano cztery lata (rys. 2) odpowiadające poszczególnym stanom cyklu koniunkturalnego na polskim rynku mieszkaniowym: okres dekonjunktury (2003 r.), początek konjunktury po wejściu do UE (2005 r.), okres hossy na rynku mieszkaniowym (2007 r.) oraz ostatni okres dekonjunktury (2009 r.).

W celu uzyskania optymalnego wektora zmiennych dyskryminacyjnych dokonano dalszej redukcji zmiennych diagnostycznych. Kolejne kroki analizy dyskryminacji wykazały istotną korelację między średnimi i wariancjami². Dokonano również przeglądu zmiennych w celu znalezienia wartości odstających. Ponieważ wartości te znajdowały się najczęściej na końcach przedziałów ze względu na każdą zmienną, przyjęto zasadę odrzucenia 10% obserwacji skrajnych, pozostawiając 80% obserwacji w kolejnych pięciu krokach. W każdej grupie znalazło się jednak pięć razy więcej obserwacji niż zmiennych dyskryminujących. Oceniono normalność wewnątrzgrupową zmiennych oraz wybrano zmienne niezależne wnoszące najbardziej istotny wkład do dyskryminacji (wykorzystano analizę krokową postę-

² Co narusza założenie analizy dyskryminacji o jednorodności wariancji/kowariancji oraz braku korelacji między średnimi i wariancjami.

pująca). W wyniku weryfikacji 19 zmiennych ostatecznie do badania pozostawiono sześć najściślej skorelowanych ze sobą i niebędących funkcją pozostałych:

- X1 zmienna grupująca (płynność wysoka, średnia, niska);
- X2 data wprowadzenia ofert do systemu (kolejne dni roku);
- X3 powierzchnia użytkowa lokalu (w m²);
- X4 cena transakcyjna mieszkania w dziesiątkach tysięcy zł/m²;
- X5 piętro w budynku, na którym znajduje się lokal;
- X6 wiek budynku w latach.



Rys. 2. Kwartalna dynamika mediany wskaźnika płynności mieszkań na rynku warszawskim w latach 2003-2009 (w dniach)

Źródło: opracowanie własne na podstawie danych z systemu MLS.

Dla tak określonego zbioru zmiennych dyskryminujących (X2-X6) oraz zmiennej dyskryminacyjnej (X1) wyznaczono współczynniki funkcji dyskryminacyjnych (tab. 1) oraz sprawdzono statystyczną istotność uzyskanych funkcji dyskryminacji. Kolejność zmiennych w ostatniej kolumnie wskazuje na siłę wpływu zmiennych (na podstawie standaryzowanych współczynników funkcji dyskryminacji). Największy wpływ na różnicowanie grup ma kwartał, w którym wprowadzana była oferta do systemu, oraz piętro, na którym znajdował się lokal.

W tabeli 2 zaprezentowano wyniki testu ch-kwadrat dla uzyskanych funkcji dyskryminacyjnych, które wskazują na istotność pierwszej funkcji w każdym badanym roku (korelacja kanoniczna bliska jedności), a wartości statystyki Lambda Wilksa bliskie zero wskazują na silną istotność mocy dyskryminującej funkcji.

Tabela 1. Podsumowanie analizy funkcji dyskryminującej dla wybranych lat 2003-2009

Lata	Lambda Wilksa	Przybliżona wartość statystyki F	Poziom p	Zmienne wybrane do przewidywania przynależności do grupy z zachowaniem kolejności wyboru
2003	0,1227	11,129	$p < 0,0000$	X2, X5, X4, X6
2005	0,2622	33,039	$p < 0,0000$	X2, X6, X5
2007	0,2202	33,931	$p < 0,0000$	X2, X4, X3
2009	0,4000	13,072	$p < 0,0000$	X2, X5

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 8.0.

Tabela 2. Wyniki testu ch-kwadrat dla kolejnych pierwiastków funkcji dyskryminacyjnych

Lata	Pierwiastek usunięty	Wartość własna	Kanoniczna R	Lambda	Chi-kwadrat	df	Poziom p
2003	0	6,571008	0,931621	0,122701	53,49904	8	0,000000
	1	0,076458	0,266509	0,928973	1,87873	3	0,597955
2005	0	2,518396	0,846038	0,262162	140,5731	6	0,000000
	1	0,084139	0,278584	0,922391	8,4826	2	0,014389
2007	0	3,457896	0,880726	0,220203	137,7019	6	0,000000
	1	0,018703	0,135498	0,981640	1,6863	2	0,430361
2009	0	1,104372	0,724430	0,400054	41,68510	4	0,000000
	1	0,187843	0,397665	0,841862	7,83232	1	0,005132

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 8.0.

Potwierdzeniem istotności pierwszej funkcji dyskryminacyjnej jest wartość własna funkcji (im wyższa, tym lepiej rozdziela grupy), natomiast skumulowany procent wariancji międzygrupowej wskazuje na wysoki udział pierwszej funkcji w dyskryminowaniu w każdym badanym roku (tab. 3).

W tabeli 4 zestawiono współczynniki funkcji klasyfikacji, opartej na liniowej kombinacji zmiennych dyskryminujących. Prawdopodobieństwo klasyfikacyjne *a priori* przyjęto proporcjonalnie do wielkości grup (ze względu na ich różną liczebność), natomiast w 2007 roku najlepsze wyniki dało prawdopodobieństwo równe. Wyznaczone wartości funkcji klasyfikacyjnych pozwalają na klasyfikację każdej transakcji do grupy o określonej płynności.

Podsumowaniem dyskryminacji są macierze klasyfikacji oparte na prawdopodobieństwie *a priori*, zawierające odsetek oraz liczbę przypadków poprawnie zaklasyfikowanych (tab. 5). Najlepiej rozdzielone zostały na trzy grupy płynności transakcje w 2003 roku (93,4% poprawnie sklasyfikowanych przypadków), najslabiej transakcje w 2007 roku (39% poprawnie sklasyfikowanych przypadków). Jednak w roku 2007 poprawnie zostały sklasyfikowane wszystkie transakcje z wysoką płynnością. Najgorsze dopasowanie uzyskano w tym roku dla transakcji o średniej płynności.

Tabela 3. Surowe współczynniki kanonicznych funkcji dyskryminacyjnych D_1 oraz D_2

Badany rok	Funkcja dyskryminacyjna	Zmienne						Wartość własna	Skumulowany procent
		X2	X3	X4	X5	X6	stała		
2003	D1	0,078		-8,040	0,191	-0,015	-2,582	6,571	0,988
	D2	0,007		-18,085	-0,335	-0,058	8,539	0,076	1,000
2005	D1	0,042			-0,020	-0,010	-3,803	2,518	0,968
	D2	0,000			-0,402	0,043	-0,028	0,084	1,000
2007	D1	0,073	0,015	2,577			-8,910	3,458	0,995
	D2	0,004	0,002	-9,372			6,957	0,019	1,000
2009	D1	-0,021			0,109		2,275	1,104	0,854
	D2	-0,007			-0,765		3,649	0,188	1,000

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 8.0.

Tabela 4. Współczynniki funkcji klasyfikacyjnej (prawdopodobieństwo klasyfikacyjne *a priori* proporcjonalne do wielkości grupy)

Badany rok	Płynność	Zmienne						p
		X2	X3	X4	X5	X6	stała	
2003	niska	0,450		167,863	4,041	0,216	-66,575	0,233
	wysoka	0,165		207,875	3,559	0,307	-54,574	0,300
	średnia	-0,029		218,340	2,892	0,314	-45,169	0,467
2005	niska	0,213			0,834	0,060	-17,439	0,505
	wysoka	0,111			1,081	0,064	-8,136	0,294
	średnia	0,054			0,791	0,112	-5,597	0,202
2007*	niska	0,876	0,402	97,583			-110,988	0,333
	wysoka	0,501	0,323	81,983			-57,361	0,333
	średnia	0,713	0,368	87,539			-81,827	0,333
2009	niska	0,092			2,500		-12,239	0,592
	wysoka	0,045			3,518		-13,405	0,102
	średnia	0,049			2,304		-7,019	0,306

* Dla 2007 roku najlepsze wyniki uzyskano przy równym podziale.

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 8.0.

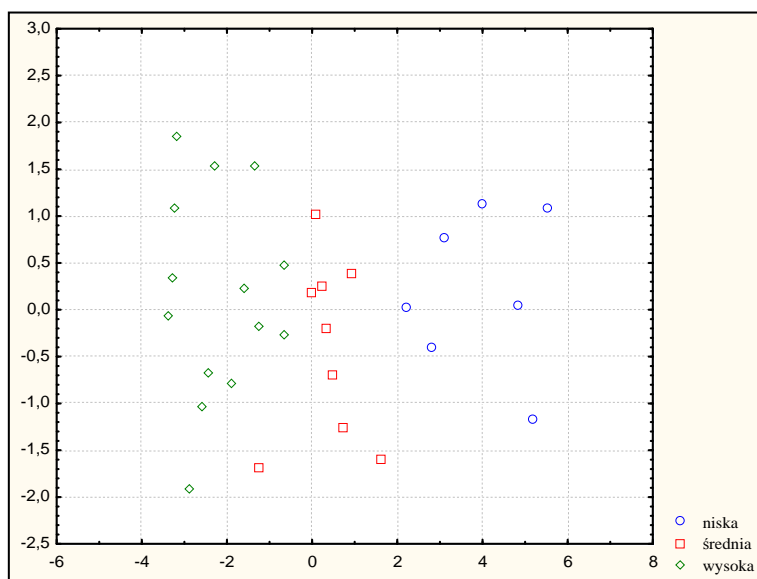
Można uznać, że transakcje klasyfikowane do grupy z wysoką płynnością we wszystkich badanych latach charakteryzują się wysokim odsetkiem poprawnej klasyfikacji.

Tabela 5. Macierz poprawności klasyfikacji

Badany rok	Płynność	niska	wysoka	średnia	p	Procent	Razem procent
2003	niska	7	0	0	0,505	100,0	93,4
	wysoka	0	8	1	0,294	88,9	
	średnia	0	1	13	0,202	92,9	
2005	niska	46	9	0	0,505	83,6	87,2
	wysoka	0	32	0	0,294	100,0	
	średnia	0	5	17	0,202	77,3	
2007	niska	12	0	28	0,333	30,0	39,0
	wysoka	0	20	0	0,333	100,0	
	średnia	0	30	5	0,333	14,3	
2009	niska	24	1	4	0,592	82,8	85,7
	wysoka	0	4	1	0,102	80,0	
	średnia	0	1	14	0,306	93,3	

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 8.0.

Na rysunku 3 przedstawiono jako przykład wykres rozrzutu dla trzech grup płynności transakcji zawartych w 2003 roku, w którym jakość klasyfikacji jest najwyższa z badanych lat. Widoczne są trzy grupy, które rozdzielone mogą być przez dwie funkcje liniowe.

**Rys. 3.** Wykres rozrzutu zmiennych kanonicznych dla transakcji zawartych 2003 roku

Źródło: opracowanie własne na podstawie danych z systemu MLS.

4. Podsumowanie

Analiza cech mieszkań wpływających na ich rynkową cenę jest przedmiotem wielu publikacji, zwłaszcza z zakresu wycen nieruchomości³. Jednak połączenie płynności rynku mieszkaniowego z fazą cyklu koniunkturalnego i oceną stabilności zestawu cech mieszkań wpływających na cenę w kolejnych fazach przeprowadzono po raz pierwszy.

W artykule wyodrębniono pięć cech mieszkań, które decydują o długości trwania oferty sprzedaży na rynku wtórnym. Weryfikowano również hipotezę, która zakładała, że na tempo sprzedaży mieszkania w systemie MLS ma wpływ zestaw tych samych cech mieszkań, niezależnie od fazy cyklu koniunkturalnego na rynku mieszkaniowym. W analizowanych latach tylko zmienna X_2 okazała się zawsze istotna ze względu na wkład do ogólnej dyskryminacji, poza nią zestaw zmiennych decydujących o przynależności do danej grupy płynności był różny w badanych latach, niezależnie od stanu rynku. W latach dekonunktury istotne okazały się zmienne X_2 oraz X_5 , nie zauważono takiej prawidłowości w latach koniunktury na rynku mieszkaniowym.

Dodatkowo sprawdzano, czy funkcja dyskryminacyjna pozwala w sposób jednoznaczny zakwalifikować mieszkania do grupy o określonej płynności. Zbudowane funkcje dyskryminacyjne w sposób zadowalający klasyfikowały mieszkania ze względu na przyjęte grupy płynności, z wyjątkiem 2007 roku, a obie funkcje dyskryminacyjne mają „niejednakowe możliwości” dyskryminowania grup.

Literatura

- Gatnar E., *Podejście wielomodalne w zagadnieniach dyskryminacji i regresji*, Wydawnictwo Naukowe PWN, Warszawa 2008.
- Gawron H., *Analiza rynku nieruchomości*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2009.
- Foryś I., *Płynność inwestycji na rynku nieruchomości na przykładzie warszawskiego rynku mieszkaniowego*, Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 612, Finanse, Rynki Finansowe, Ubezpieczenia nr 28, Szczecin 2010.
- Kucharska-Stasiak E., *Nieruchomość w gospodarce rynkowej*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Stanisz A., *Przystępny kurs statystyki z zastosowaniem Statistica Pl na przykładach z medycyny*, t. III, *Analizy wielowymiarowe*, Wydawnictwo StatSoft Polska Sp. z o.o., Kraków 2007.

³ Przykładem są liczne artykuły zamieszczone w zeszytach naukowych Towarzystwa Naukowego Nieruchomości, kwartalniku „Świat Nieruchomości” (wyd. Fundacja Uniwersytetu Ekonomicznego w Krakowie) oraz wcześniejszych monografiach Sekcji Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego.

**MULTIDIMENSIONAL ANALYSIS OF FEATURES
OF APARTMENTS SOLD ON WARSAW REAL ESTATE MARKET
IN THE RESEARCH OF DURATION OFFER IN MLS SYSTEM**

Summary: The subject of the analysis is the Warsaw home market. The hypothesis on the stability of features influencing liquidity, irrespective of a situation on the estate market, is negative. For this purpose the author uses the discrimination function. It is shown that the discrimination functions in a satisfactory way classify transactions into three proposed liquidity groups: low, high and average.