

**PODZIAŁ POLSKI NA OBSZARY ZE SZCZEGÓLNYM
UWZGLĘDNIENIEM PODREGIONÓW ZAGROŻONYCH
UBÓSTWEM**

Jadwiga Zaród

*Zachodniopomorski Uniwersytet Technologiczny w
Szczecinie*

Katedra Zastosowań Matematyki w Ekonomii

Cel pracy

Celem tej pracy jest ukazanie zróżnicowania rozwoju podregionów w Polsce ze względu na wybrane cechy (warunki ekonomiczne, społeczne i mieszkaniowe ludności oraz jej aktywność zawodowa) i charakterystyka obszarów zagrożonych ubóstwem.

Material badawczy

Badaniami objęto 66 podregionów Polski, które opisano za pomocą zmiennych diagnostycznych:

- x_1 – ludność korzystająca z oczyszczalni ścieków w % ludności ogółem,
- x_2 – redukcja zanieczyszczeń powietrza w % zanieczyszczeń wytworzonych,
- x_3 – odpady komunalne zebrane na 1 mieszkańca w kg,
- x_4 – liczba ludności na 1 pracującego,
- x_5 – bezrobotni zarejestrowani ogółem w tys. osób,
- x_6 – stopa bezrobocia w %,
- x_7 – przeciętne miesięczne wynagrodzenie brutto w zł,
- x_8 – sieć rozdzielcza wodociągowa w km/100km²,
- x_9 – sieć rozdzielcza kanalizacyjna w km/100km²,
- x_{10} – sieć rozdzielcza gazowa w km/100km²,
- x_{11} – przeciętna powierzchnia użytkowa mieszkania w m²,
- x_{12} – przeciętna powierzchnia użytkowa na 1 mieszkańca w m²,
- x_{13} – liczba ludności na 1 łóżko w szpitalach ogólnych,
- x_{14} – liczba czytelników bibliotek na 1000 ludności
- x_{15} – liczba uczniów liceów,
- x_{16} – liczba uczniów techników,
- x_{17} – liczba studentów,
- x_{18} – miejsca na widowni w kinach na 1000 ludności,
- X_{19} - produkcja sprzedana przemysłu na 1mieszkańca w zł
- x_{20} – liczba podmiotów gospodarki narodowej wpisanych do rejestru REGON,
- X_{21} – liczba podmiotów gospodarki narodowej na 10 tys. ludności.

Materiał badawczy

Na podstawie wartości zmiennych diagnostycznych dokonano wstępnego podziału podregionów na 7 obszarów (grup). Liczbę grup k ustalono na podstawie wzoru:

$$k = 1 + 3,322 \log n \quad (1)$$

gdzie $n = 66$ – liczba podregionów.

Dane wejściowe poddano standaryzacji, aby uniezależnić wyniki analiz od jednostek pomiaru poszczególnych zmiennych za pomocą wzoru:

gdzie:

$$Z_{ij} = \frac{x_{ij} - \bar{x}}{s} \quad (2)$$

x_{ij} - wartość j -tej zmiennej dla i -tego podregionu,

\bar{x} - wartość średnia danej zmiennej,

s – odchylenie standardowe.

Zmiennae x_4 , x_5 i x_6 są destymulantą, poprzez przemnożenie ich wartości przez minus jeden dokonano przekształcenia w stymulantę.

Material badawczy

Analiza dyskryminacyjna zakłada, że dane (ujęte w postaci zmiennych) reprezentują próbę z wielowymiarowego rozkładu normalnego. Po wykonaniu histogramów rozkładów liczebności oraz testów χ^2 i Kołmogorowa-Smirnowa, służących do oceny normalności rozkładu zredukowano liczbę zmiennych do: $x_1, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{21}$. Dalsze ograniczenie listy zmiennych spowodowało badanie ich mocy dyskryminacyjnej za pomocą testu Lambda Wilksa i testu F.

Materiał badawczy

Tabela 1. Wyniki testu Lambda Wilksa i testu F

Zmienne	Lambda Wilksa	F	P	Tolerancja	1-Tolerancja (R kwadrat)
x_1	0,0095	7,4420	0,0000	0,4201	0,5799
x_4	0,0108	9,7070	0,0000	0,3446	0,6554
x_6	0,0146	16,2198	0,0000	0,2937	0,7063
x_7	0,0065	2,3472	0,0439	0,9113	0,0887
x_{11}	0,0097	7,8214	0,0000	0,3710	0,6290
x_{12}	0,0077	4,4510	0,0010	0,6003	0,3997
x_{21}	0,0075	4,1156	0,0018	0,4666	0,5334

Metody badawcze

Analiza dyskryminacyjna, główna metoda badawcza, stanowi zespół metod dyskryminacyjnych i klasyfikacyjnych.

Metody dyskryminacyjne podziela Polskę na obszary jak najbardziej do siebie podobne ze względu na opisujące je cechy.

Metody klasyfikacyjne ustala, do którego z utworzonych obszarów należy przyporządkować dany podregion, wykorzystując w tym celu te zmienne, które miały największą moc dyskryminacyjną

Metody dyskryminacyjne

Do podziału wykorzystano funkcje dyskryminacyjne o postaci:

$$F = A^T Z$$

gdzie:

F – macierz zmiennych dyskryminacyjnych

A^T – transponowana macierz współczynników zmiennych dyskryminacyjnych

Z – wystandaryzowana macierz obserwacji

Maksymalna liczba obliczonych funkcji jest równa liczbie grup minus jeden (czyli 6).

Funkcja dyskryminacyjna o postaci:

$$F = 0,4910x_1 + 0,0776x_4 + 1,1047x_6 + 0,3394x_7 - 0,7740x_{11} + 0,0967x_{12} - 0,2289x_{21}$$

wyjaśnia 70,2% wariancji międzygrupowej, ma najniższą wartość testu Lambda Wilksa (0,0051) i dlatego będzie stanowiła podstawę dalszych analiz.

Metody dyskryminacyjne

Na podstawie funkcji dyskryminacyjnej obliczono wartości kanoniczne dla wszystkich podregionów i przeciętne dla każdego obszaru.

Tabela 2. Wartości średnie funkcji dyskryminacyjnej obszarów

Obszary	Średnie wartości kanoniczne
1	-3,9058
2	-1,6015
3	-1,5640
4	-0,9360
5	1,5449
6	4,3921
7	8,1063

Najniższa przeciętna wartość funkcji dyskryminacyjnej wskazuje obszar najslabiej rozwinięty pod względem badanych cech. Podregiony należące do tego obszaru są zagrożone ubóstwem i wykluczeniem społecznym.

Metody klasyfikacyjne

Do klasyfikacji obiektów wykorzystano funkcje klasyfikacyjne o postaci:

$$K_r = c_r + c_{r1}x_1 + c_{r2}x_2 + \dots + c_{rj}x_j \quad r = 1, 2, \dots, k, \quad j = 1, 2, \dots, m$$

gdzie:

K_r – wartość r -tej zmiennej klasyfikacyjnej (dla r -tego obszaru)

c_{r0} – stała dla r -tego obszaru

c_{rj} – współczynniki zmiennych o istotnej mocy dyskryminacyjnej

x_j – wartości obserwowane (standaryzowane) dla j -tej zmiennej.

Funkcje o takiej postaci zostały oszacowane dla 7 obszarów. Dla każdego podregionu obliczono wartości wszystkich funkcji klasyfikacyjnych. Dany podregion przyporządkowano do obszaru, dla którego ma on największą wartość klasyfikacyjną.

Metody klasyfikacyjne

Tabela 3. Klasyfikacja wstępna podregionów

Obszar	Poprawność klasyfikacji (%)	Liczba podregionów w poszczególnych obszarach						
		I	II	III	IV	V	VI	VII
I	87,50	7	1					
II	83,33		10	2				
III	86,67		2	13				
IV	63,64			2	7	2		
V	62,50				3	5		
VI	83,33					1	5	
VII	100,00							6
średnio	80,30	7	13	17	10	8	5	6

wiersze – klasyfikacja wstępna,

kolumny – podział na podstawie funkcji klasyfikacyjnych

Metody klasyfikacyjne

Tabela 4. Współczynniki funkcji klasyfikacyjnych

Obszary Zmienne	I	II	III	IV	V	VI	VII
x_1	-0,5673	4,0386	-3,3322	-4,2701	1,5410	6,1751	6,1698
x_4	3,5741	6,1451	-1,4459	-6,4698	-3,2295	2,0526	2,9246
x_6	-15,1079	-10,2765	-1,1845	2,8683	7,1011	9,1054	15,5272
x_7	5,5132	5,1421	0,2519	2,5813	-2,7208	-5,2179	-8,9572
x_{11}	-3,2786	-3,9356	0,7139	3,3536	2,4103	-0,8036	0,1772
x_{12}	2,9953	0,2264	0,3409	2,3751	-3,4598	-3,3156	-0,6356
x_{21}	-1,5134	-0,3857	-1,6596	0,6934	1,1499	2,0437	5,0008
stała	-13,5819	-8,1303	-2,8016	-9,7944	-5,0367	-13,3517	-36,1211

Metody klasyfikacyjne

Tabela 5. Wyniki klasyfikacji końcowej

Obszary	Poprawność klasyfikacji (%)	Liczba podregionów w poszczególnych obszarach						
		I	II	III	IV	V	VI	VII
I	100	7						
II	100		8					
III	100			25				
IV	100				7			
V	100					7		
VI	100						6	
VII	100							6
średnio	100	7	8	25	7	7	6	6

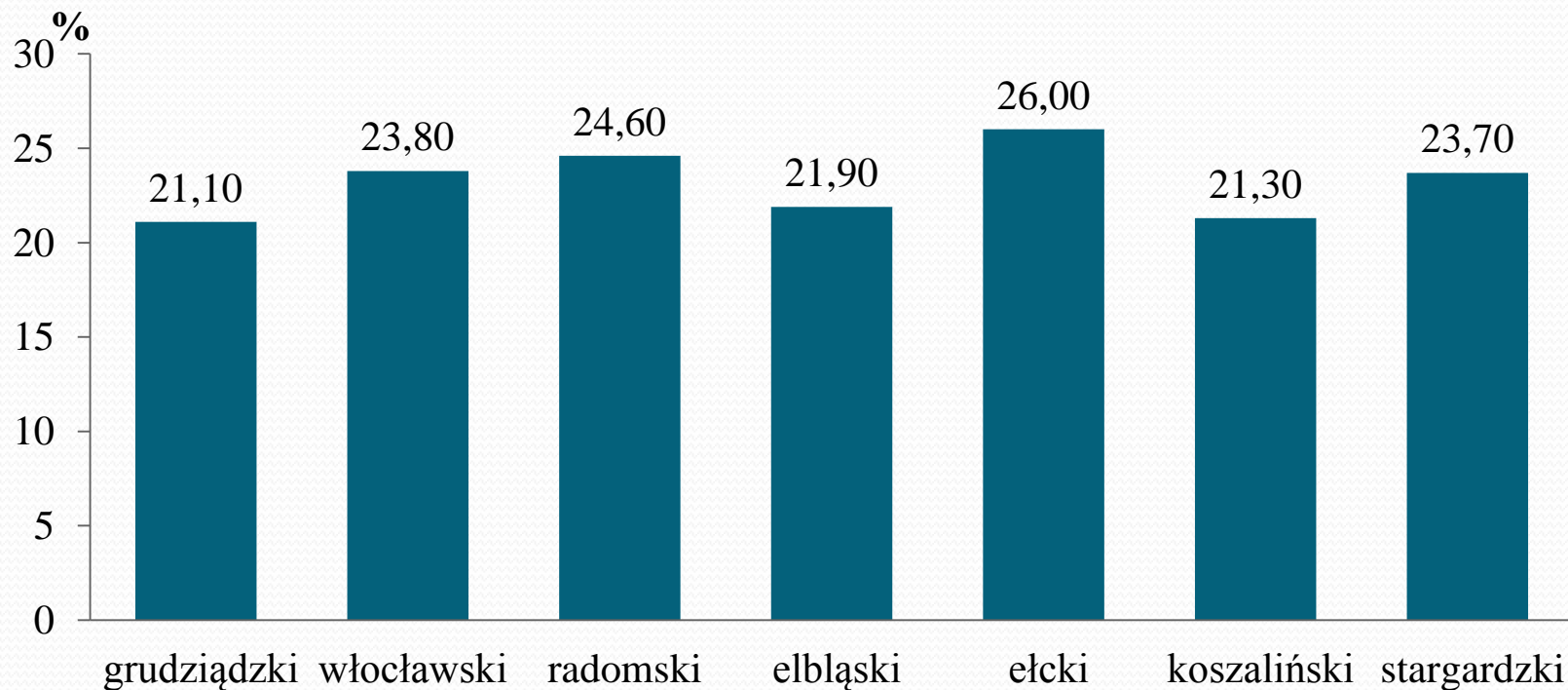
Klasyfikacja podregionów do odpowiednich obszarów

Obszar	Jednostki terytorialne
I	Podregion grudziądzki, podregion włocławski, podregion radomski, podregion elbląski, podregion ełcki, podregion koszaliński, podregion stargardzki.
II	Podregion krośnieński , podregion przemyski, podregion rzeszowski, podregion tarnobrzeski, podregion słupecki, podregion starogardzki, podregion kielecki, podregion olsztyński.
III	Podregion jeleniogórski, podregion wałbrzyski , podregion bielski, podregion chełmsko-zamojski, podregion puławski, podregion gorzowski, podregion zielonogórski, podregion łódzki, podregion piotrkowski, podregion sieradzki, podregion skierniewicki, podregion nowosądecki, podregion oświęcimski, podregion tarnowski, podregion ciechanowsko-płocki, podregion ostrołęcko-siedlecki, podregion nyski, podregion łomżyński, podregion suwalski, podregion częstochowski, podregion sandomiersko-jędrzejowski, podregion kaliski, podregion koniński, podregion leszczyński, podregion pilski.

Klasyfikacja podregionów do odpowiednich obszarów

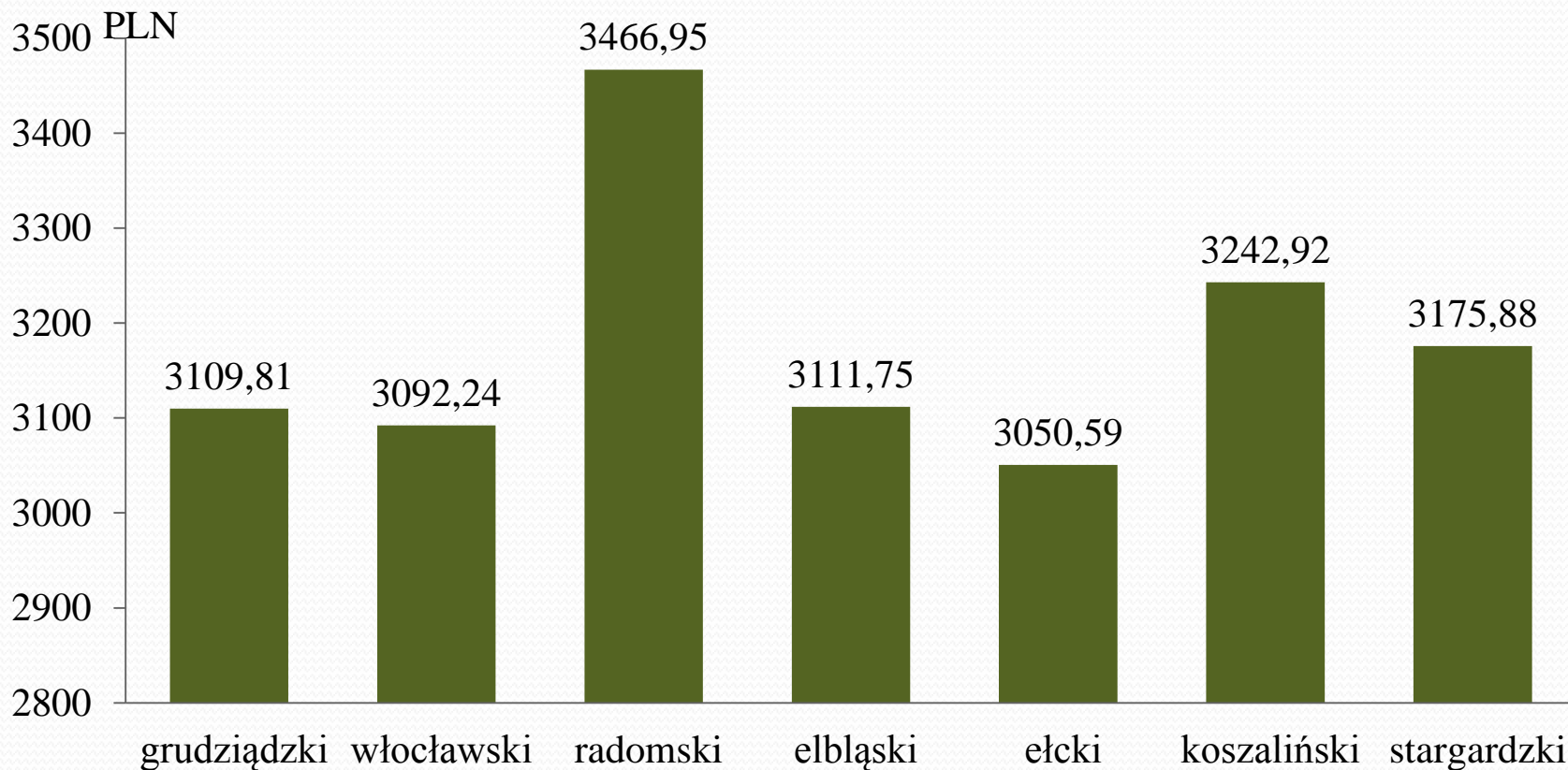
Obszar	Jednostki terytorialne
IV	Podregion wrocławski, podregion krakowski, podregion warszawski wschodni, podregion warszawski zachodni, podregion gdański, podregion poznański, podregion szczeciński.
V	Podregion lubelski, podregion opolski, podregion białostocki, podregion bielski, podregion bytomski, podregion rybnicki, podregion sosnowiecki.
VI	Podregion legnicko-głogowski, podregion bydgosko-toruński, miasto łódź, podregion gliwicki, podregion tyski, miasto Szczecin.
VII	Miasto Wrocław, miasto Kraków, miasto Warszawa, podregion trójmiejski, podregion katowicki, miasto Poznań.

Charakterystyka podregionów zagrożonych ubóstwem



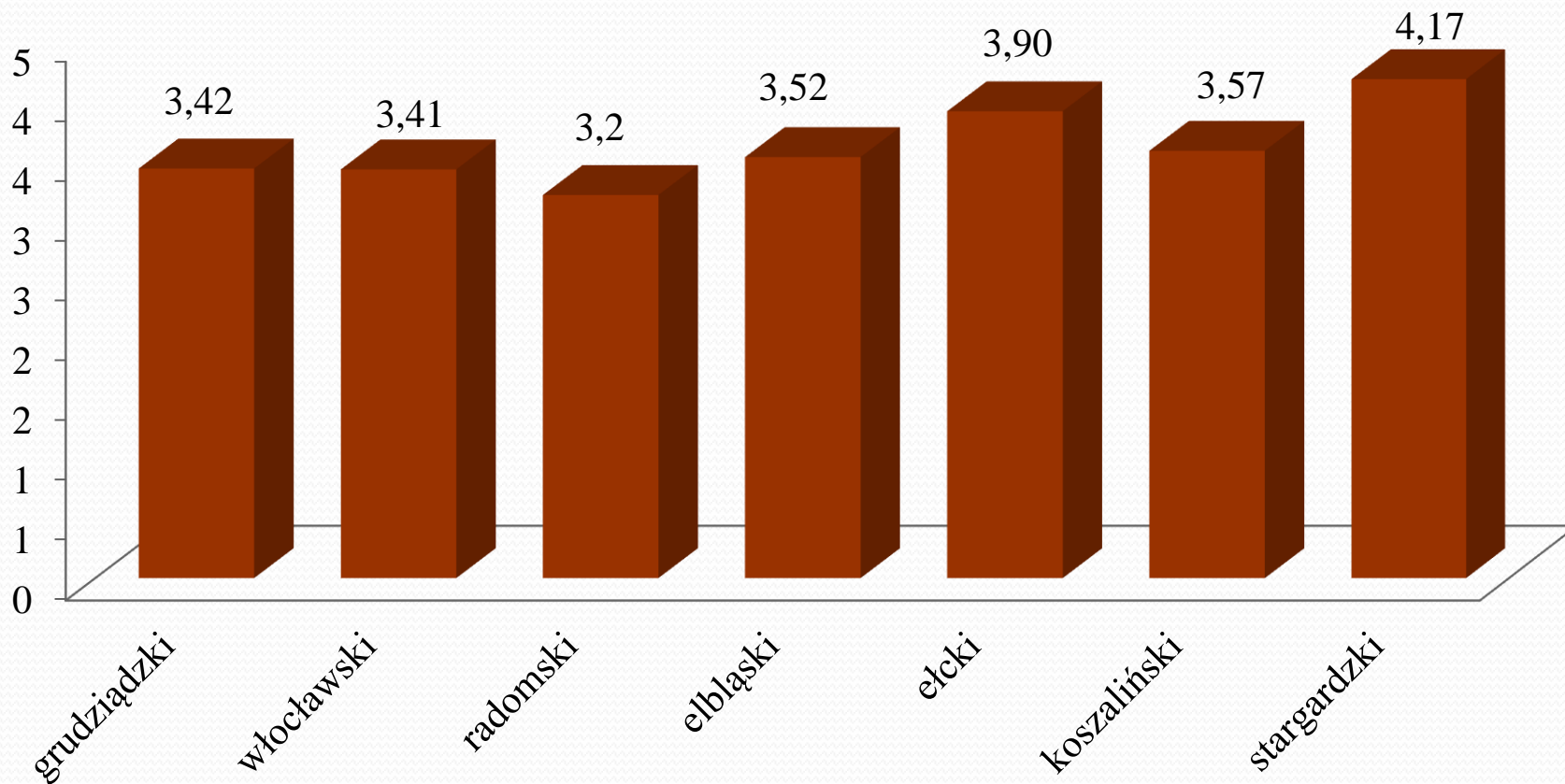
Wykres 1. Stopa bezrobocia (%) podregionów I obszaru

Charakterystyka obszaru I



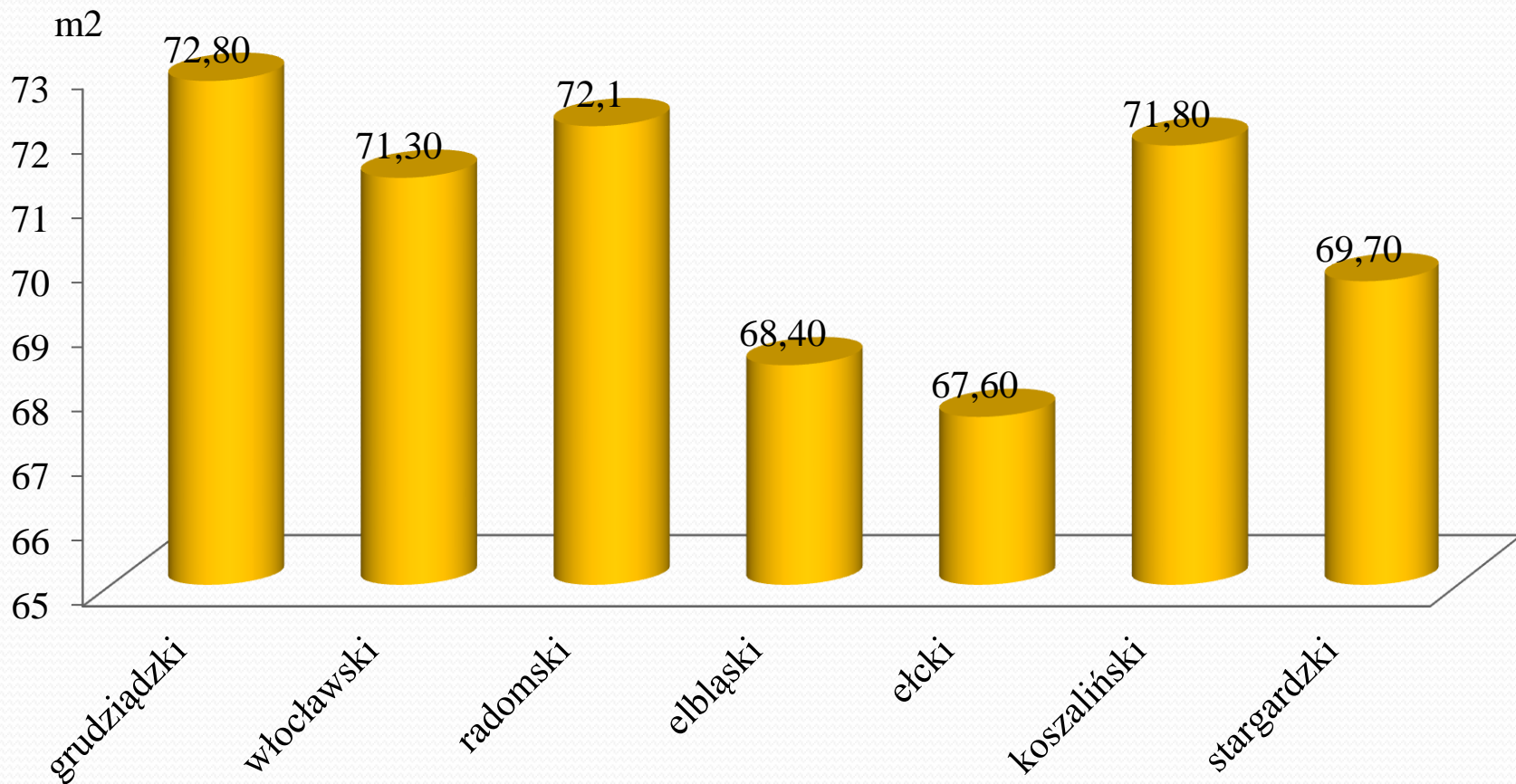
Wykres 2. Przeciętne miesięczne wynagrodzenie brutto w zł

Charakterystyka obszaru I



Wykres 3. Liczba ludności przypadająca na 1 zatrudnionego

Charakterystyka obszaru I



Wykres 4. Przeciętna powierzchnia użytkowa mieszkania

Wnioski

1. Funkcje dyskryminacyjne pozwoliły podzielić Polskę na obszary jak najbardziej do siebie podobne ze względu na wybrane zmienne.
Trafność podziału potwierdza niski współczynnik testu Lambda Wilksa (0,0051)
2. Wyznaczone wartości funkcji dyskryminacyjnej wskazują podregiony zagrożone ubóstwem i wykluczeniem społecznym i uszeregowują wyłonione kompleksy od najslabiej do najlepiej rozwiniętych.

Wnioski

3. Funkcje klasyfikacyjne pozwalają przydzielić poszczególne podregiony do wyznaczonych obszarów.
4. Analiza dyskryminacyjna może być wykorzystywana jako narzędzie wspomagające badanie zróżnicowania poziomu życia społeczeństwa.



Dziękuję