

Wpływ obserwacji odstających na wartości indeksu cen mieszkań w Polsce

Ośrodek Statystyki Rynku Nieruchomości

Iwona Foryś
Mariusz Doszyń

Etapy badania

1. Metody wyznaczania HPI
2. Identyfikacja outliersów
3. HPI w 2 wariantach – bez odrzucania i z odrzucaniem outliersów
4. Porównanie indeksów
5. Wnioski i dalsze kierunki badań

1. Metody wyznaczania HPI

Metoda stratyfikacji

Zdefiniowanie podgrup, dla których będą liczone ważone HPI

Wagi – wartość transakcji lokalami mieszkalnymi w 2023 roku (w danej podgrupie)

Baza danych – LOKALE

Zmienne grupujące:

- województwa
- miasta (duże miasta wojewódzkie, pozostałe miasta)
- dzielnice (wybrane dzielnice Warszawy)
- liczba izb
- rodzaj rynku (pierwotny, wtórny)

1. Metodyka wyznaczania HPI

Przykładowe podgrupy (straty) dla wybranego województwa:

Województwo	miasta	liczba izb	rodzaj rynku	liczba obserwacji/ 2023/4	średnia cena/ 2023/4	liczba obserwacji/ 2024/1	średnia cena/ 2024/1	HPI	waga
Dolnośląskie	pozostałe	1 lub 2	pierwotny	317	7648,2	290	8005,7	1,047	423 303 297
Dolnośląskie	pozostałe	1 lub 2	wtórny	643	5702,5	771	6143,2	1,077	491 838 640
Dolnośląskie	pozostałe	3	pierwotny	438	7046,3	377	7594,8	1,078	648 410 917
Dolnośląskie	pozostałe	3	wtórny	1201	5423,3	1388	5541,1	1,022	1 042 412 991
Dolnośląskie	pozostałe	>3	pierwotny	359	6319,2	370	6704,6	1,061	810 828 238
Dolnośląskie	pozostałe	>3	wtórny	890	5033,3	935	5267,3	1,046	942 763 593
Dolnośląskie	Wrocław	1 lub 2	pierwotny	584	11695,7	527	11481,6	0,982	1 516 888 063
Dolnośląskie	Wrocław	1 lub 2	wtórny	503	12624,0	606	13280,3	1,052	873 708 018
Dolnośląskie	Wrocław	3	pierwotny	397	10859,6	447	10769,9	0,992	1 306 953 655
Dolnośląskie	Wrocław	3	wtórny	684	10889,6	686	11287,2	1,037	1 332 564 740
Dolnośląskie	Wrocław	>3	pierwotny	152	10189,1	227	9507,2	0,933	1 003 677 081
Dolnośląskie	Wrocław	>3	wtórny	392	9707,1	389	10085,9	1,039	866 811 639

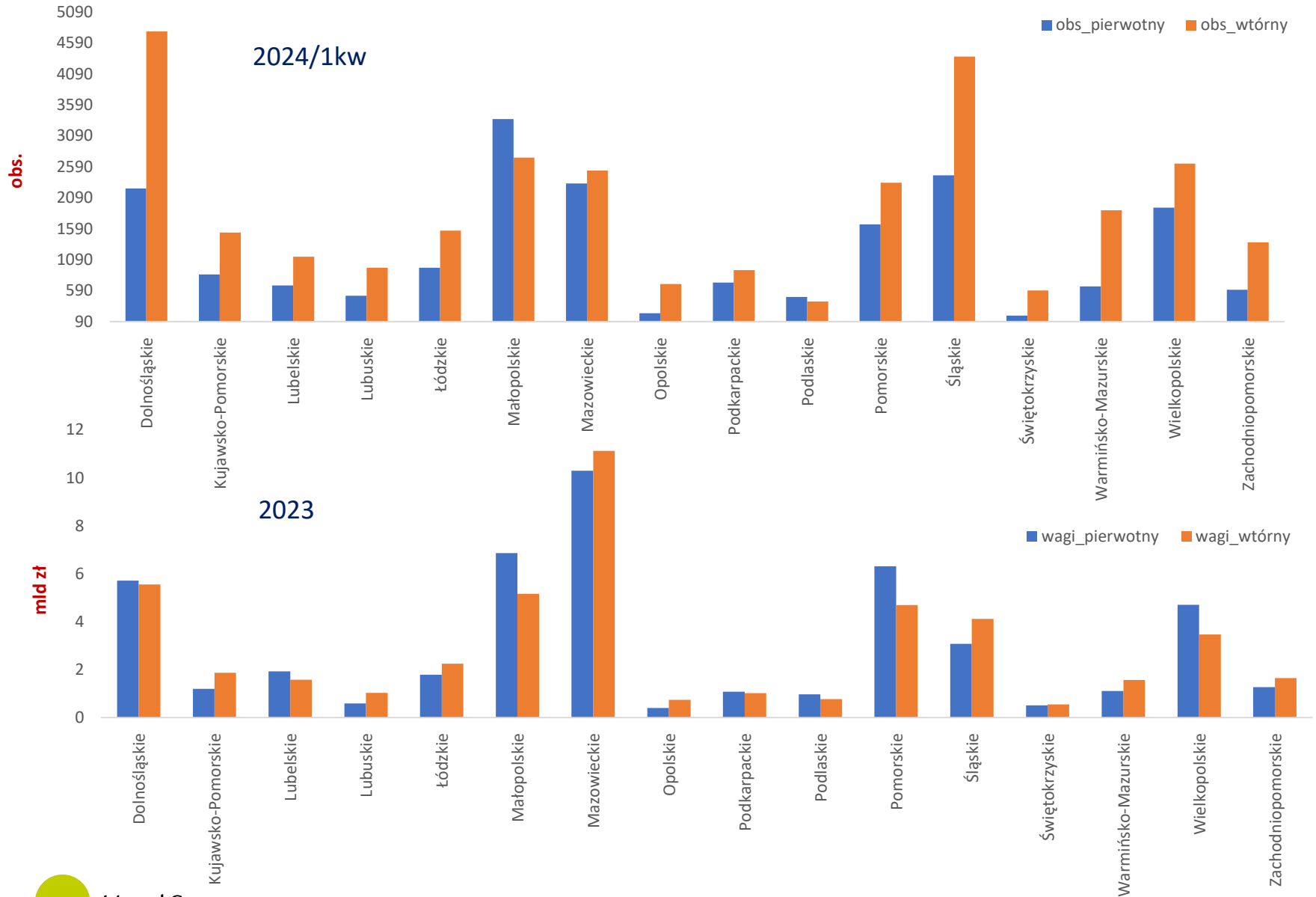
1. Metody wyznaczania HPI

Surowy wskaźnik cen nieruchomości będzie w kolejnych okresach zniekształcany z uwagi na różnice w jakości nieruchomości będących przedmiotem obrotu w kolejnych okresach

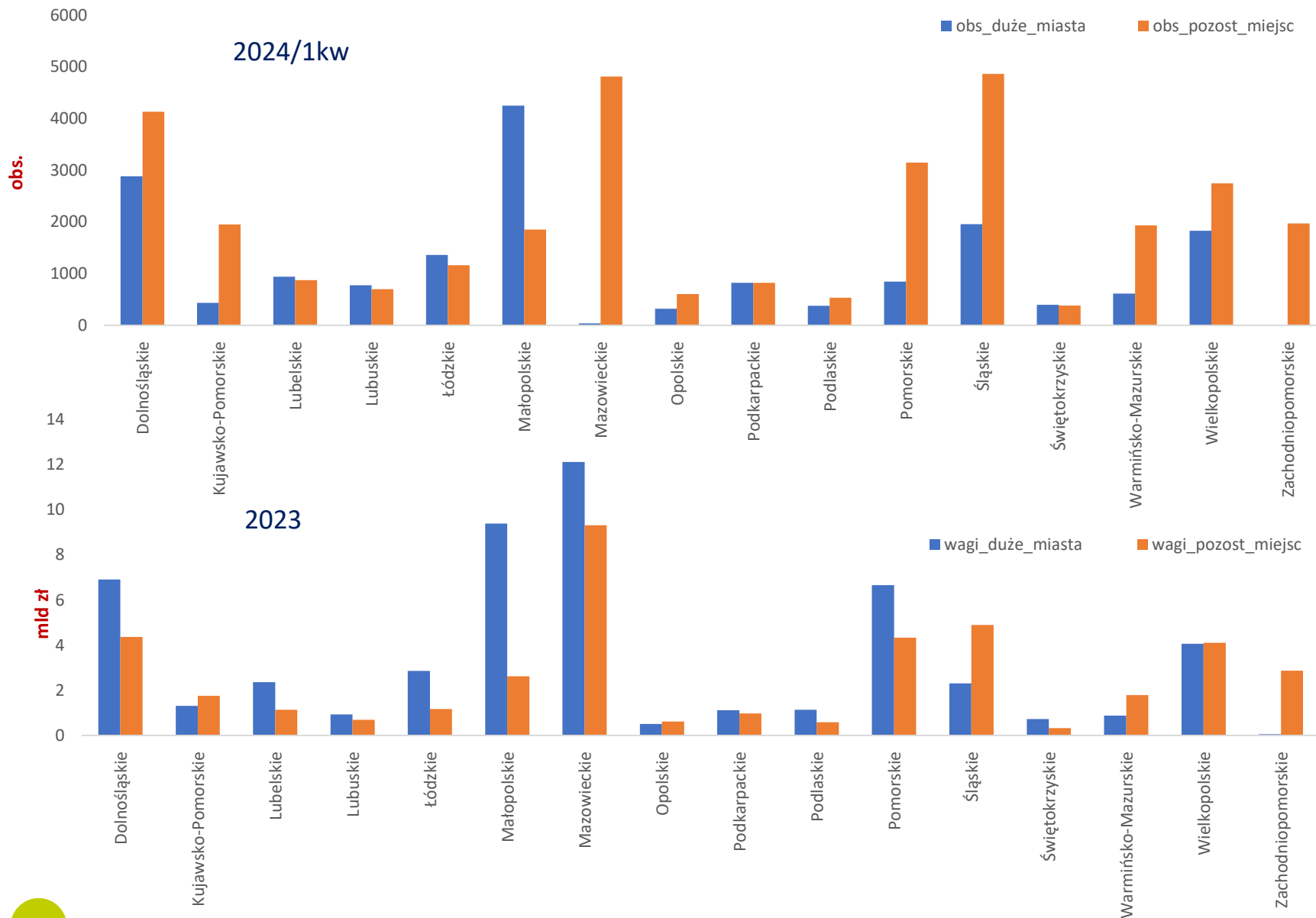
Z tego powodu najczęściej w praktyce wykorzystuje się sposoby konstruowania indeksów cen skorygowanych cechami nieruchomości:

- stratyfikacja (średnie lub mediany)
- indeksy powtórnej sprzedaży
- indeksy hedoniczne

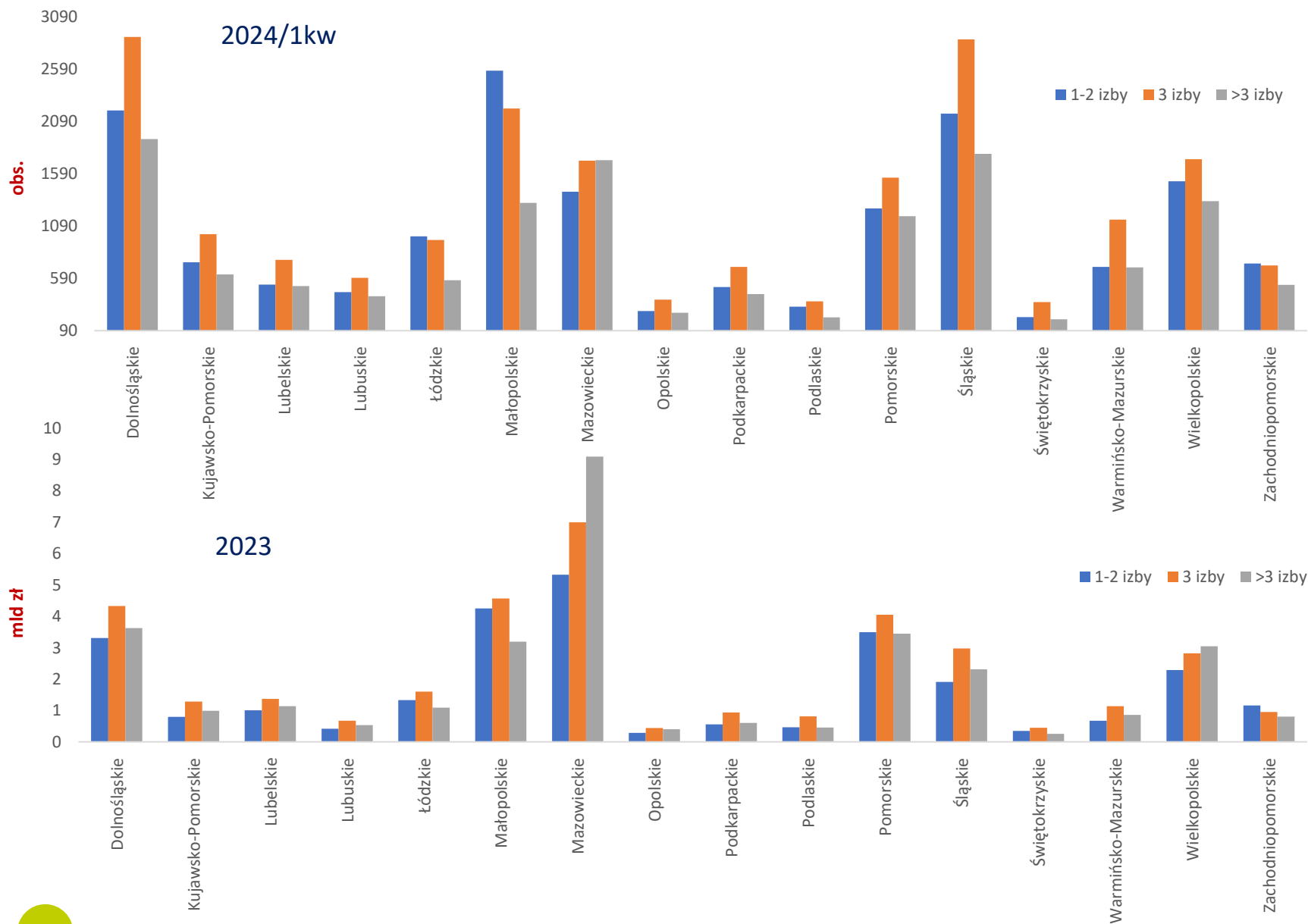
Liczba obserwacji (24/1) i wagi (2023) w województwach – rynek pierwotny i wtórny



Liczba obserwacji (24/1) i wagi (2023) w województwach – miasta duże i pozostałe



Liczba obserwacji (2024/1) i wagi (2023) w województwach – liczba izb



2. Identyfikacja outliersów

Metoda Tukey'a

outliers = 1 jeżeli:

$\text{cena } 1\text{m}^2 < Q_1 - 1,5 * \text{IQR}$

lub

$\text{cena } 1\text{m}^2 > Q_3 + 1,5 * \text{IQR}$

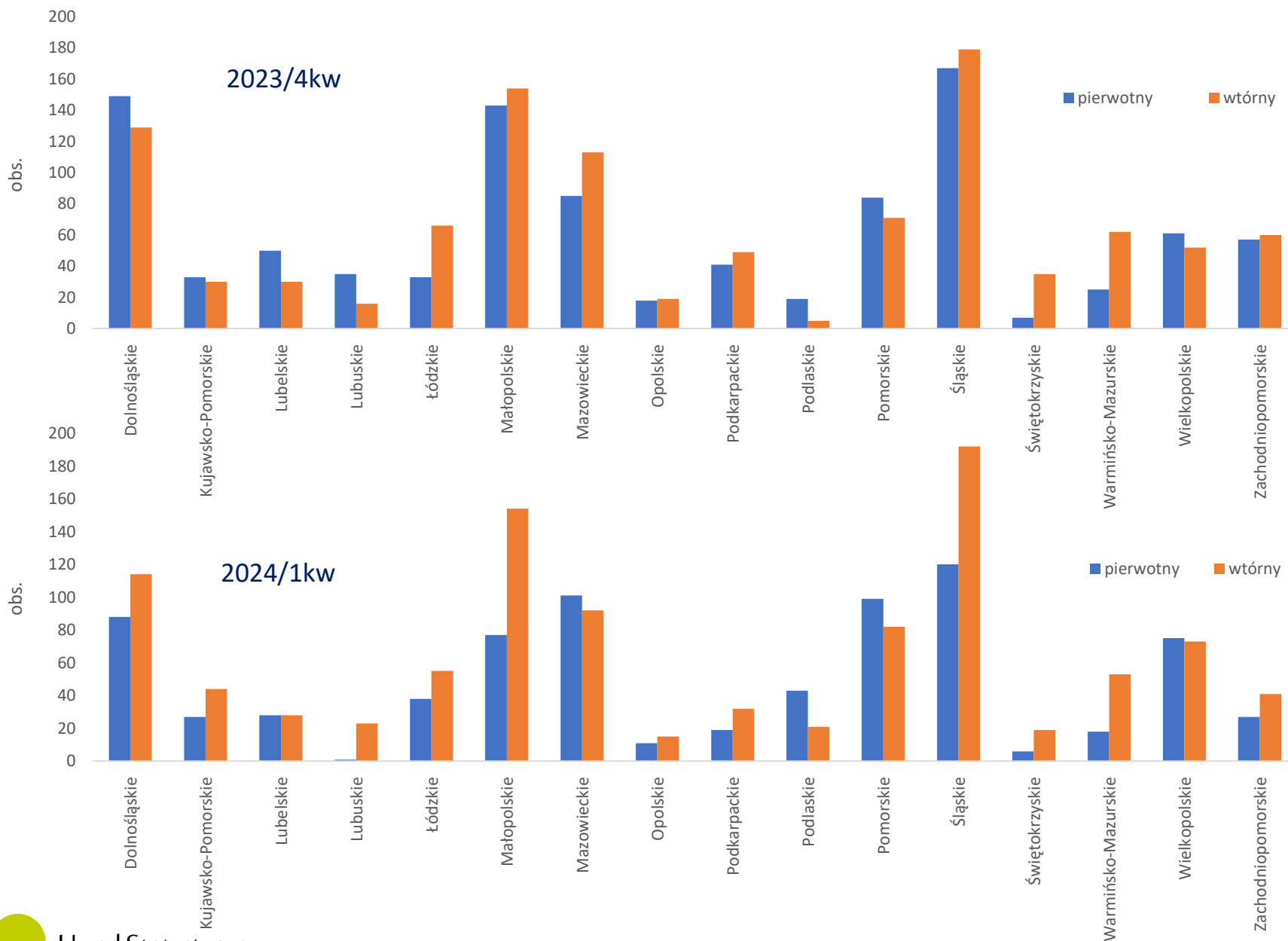
Q_1, Q_3 – kwartył 1 i 3

$\text{IQR} = Q_3 - Q_1$

Zmienne grupujące:

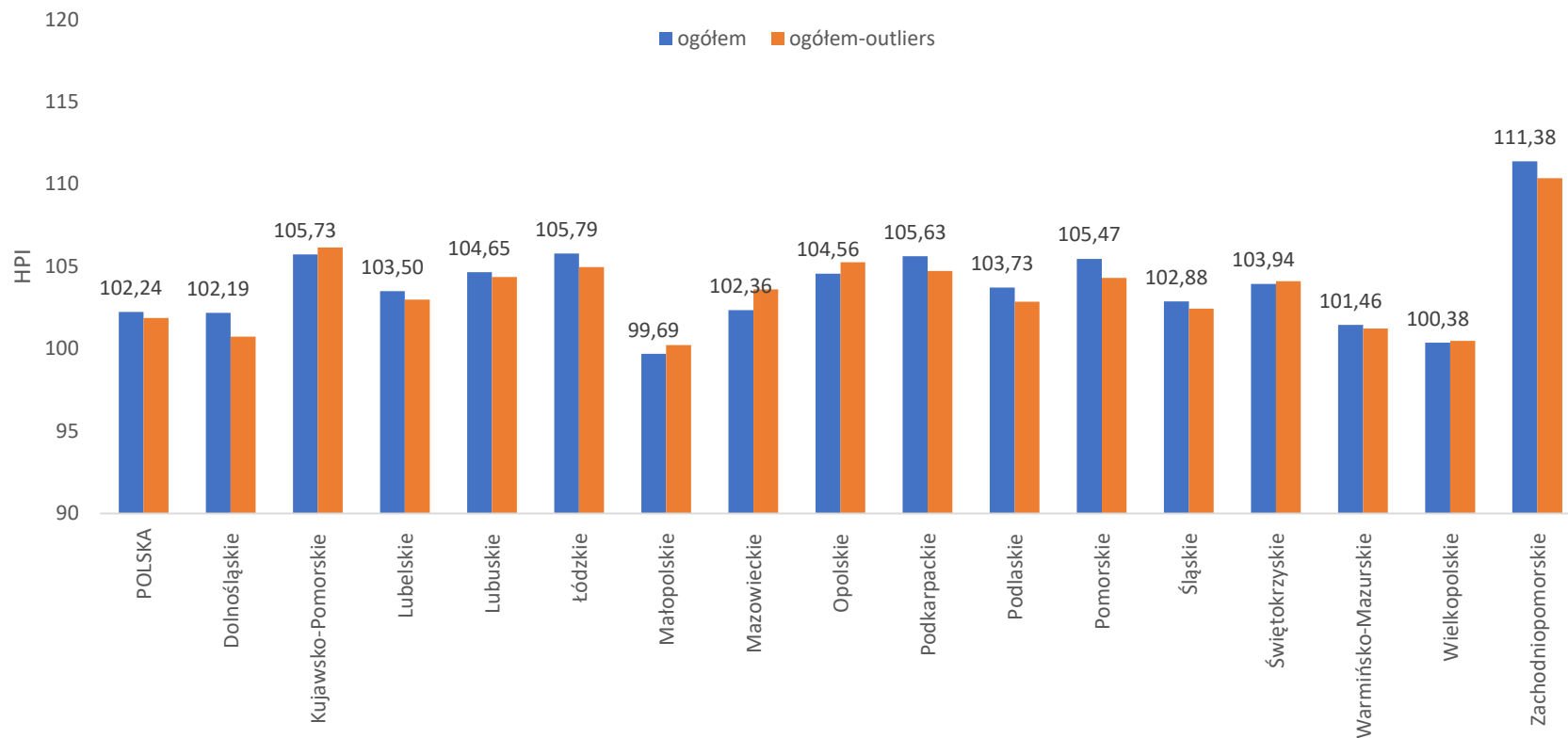
- województwa
- powiaty
- rodzaj rynku (pierwotny, wtórny)

Liczba outliersów w 2023/4kw oraz 2024/1kw w województwach



3. HPI w 2 wariantach

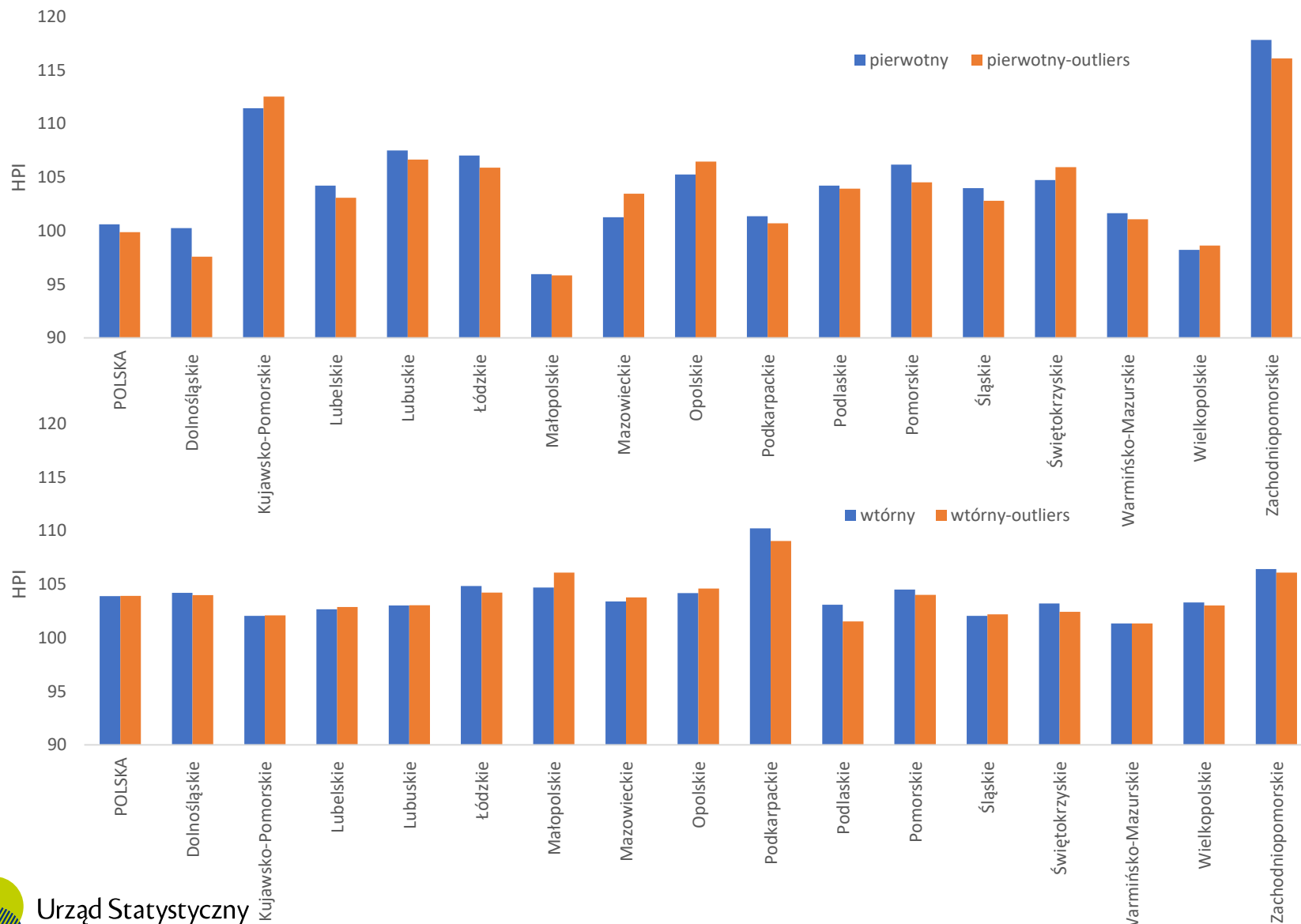
HPI dla Polski i województw ogółem (2024/1kw)



Różnice między HPI i HPI po wyeliminowaniu outliersów

Rynek	ME [pp]	MAE [pp]
ogółem	0,29	0,66
pierwotny	0,39	1,10
wtórny	0,17	0,48

Wpływ outliersów na HPI dla Polski i województw, rynek pierwotny i wtórny



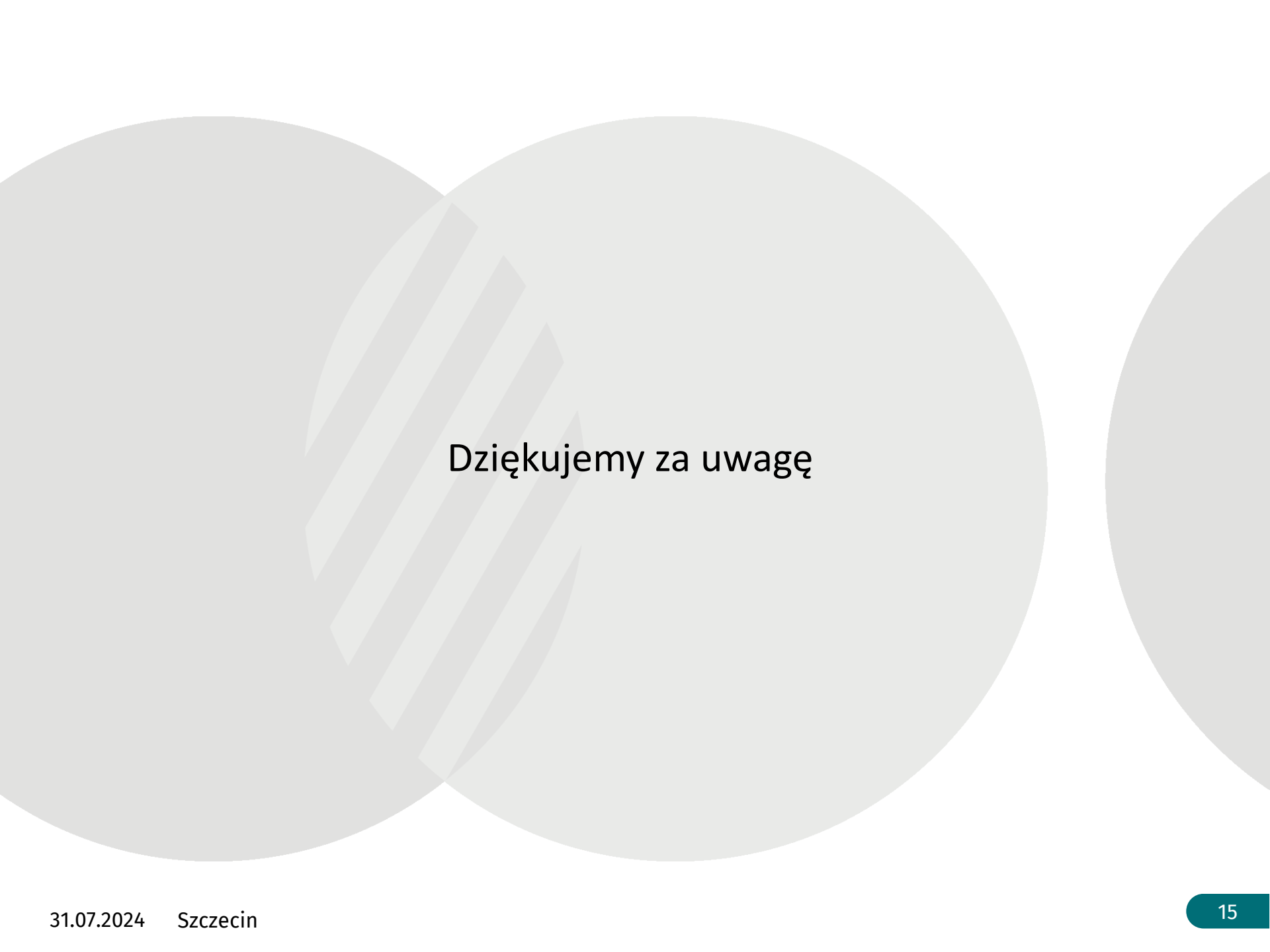
4. Różnice między HPI i HPI po wyeliminowaniu outliersów (2024/1kw, 2023/4kw = 100)

HPI - HPI[-outliers]	ogółem	pierwotny	wtórny
POLSKA	0,36	0,72	-0,02
Dolnośląskie	1,46	2,67	0,20
Kujawsko-Pomorskie	-0,44	-1,08	-0,03
Lubelskie	0,51	1,12	-0,21
Lubuskie	0,29	0,86	-0,02
Łódzkie	0,82	1,12	0,59
Małopolskie	-0,53	0,11	-1,39
Mazowieckie	-1,26	-2,20	-0,38
Opolskie	-0,69	-1,19	-0,43
Podkarpackie	0,90	0,65	1,18
Podlaskie	0,87	0,31	1,58
Pomorskie	1,17	1,65	0,50
Śląskie	0,43	1,20	-0,14
Świętokrzyskie	-0,16	-1,20	0,80
Warmińsko-Mazurskie	0,23	0,55	0,00
Wielkopolskie	-0,11	-0,40	0,29
Zachodniopomorskie	1,01	1,73	0,35

5. Wnioski i dalsze kierunki badań

Wnioski

- odrzycanie obserwacji odstających (outliersów) wpływa na HPI
- budowa hedonicznych indeksów cen (z/bez cen odstających)
- poszukiwanie innych/ dodatkowych źródeł danych niż RCN



Dziękujemy za uwagę