



AGH

AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

Projekt 1

Ekonometria Przestrzenna

Maciej Nagły, Karol Kuciński, Krzysztof Księżyc

Informatyka i Ekonometria

II stopień, studia stacjonarne

Spis treści

Wstęp.....	3
Dane.....	3
Globalne miary autokorelacji przestrzennej.....	4
Stystyka <i>I</i> Morana.....	4
Statystyka <i>C</i> Geary'ego.....	8
Macierz sąsiedztwa rzędu I.....	9
Macierz sąsiedztwa rzędu IV.....	10
Macierz sąsiedztwa oparta na centroidach.....	11
Statystyka <i>G</i> Getisa i Orda.....	11
Macierz sąsiedztwa rzędu I.....	12
Macierz sąsiedztwa rzędu IV.....	13
Macierz sąsiedztwa oparta na centroidach.....	14
Korelogram i wykres rozproszenia.....	15
Wspólna Interpretacja.....	17
Podsumowanie.....	18

Wstęp

W analizie została przeprowadzona ocena autokorelacji przestrzennej średniej ceny mieszkań za metr kwadratowy.

W badaniu wykorzystano kluczowe miary autokorelacji przestrzennej takie jak statystyka I Morana, C Geary'ego, G Getisa oraz Orda.

Istotność statystyk została oceniona na podstawie testu permutacyjnego oraz testu opartego na statystyce Z.

Miary te pozwalają na ocenę, czy analizowana zmienna wykazuje tendencję do podobnych wartości w sąsiadujących jednostkach przestrzennych, czy też charakteryzuje się bardziej losowym rozkładem przestrzennym.

Celem analizy jest nie tylko określenie istnienia autokorelacji przestrzennej, ale także ocena jej istotności statystycznej. W związku z tym przeprowadzono różne warianty testów istotności, co pozwoliło na uzyskanie bardziej wiarygodnych i precyzyjnych wniosków. Uwzględnione zostały również różne formy macierzy wag przestrzennych, w tym macierze sąsiedztwa oraz macierze odległości, które umożliwią dokładniejsze uchwycenie zależności przestrzennych między jednostkami.

Dane

Do analizy zostały wykorzystane mapki z podziałem na województwa pobrane z poniższej strony GIS'u:

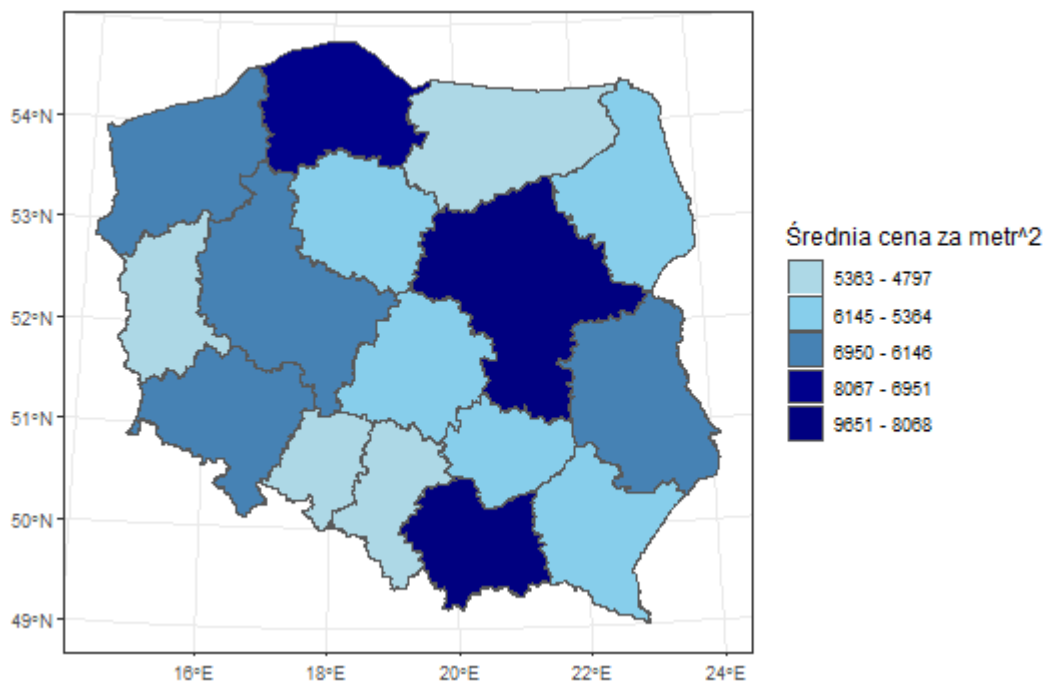
[GIS Support Granice administracyjne - GIS Support](#)

Zbiór danych zawierający średnią cenę mieszkań za metr kwadratowy dla każdego województwa na rok 2022 został pobrany z oficjalnej strony GUS'u:

[GUS - Bank Danych Lokalnych](#)

JPT_NAZWA_	ogółem
dolnośląskie	6928
kujawsko-pomorskie	5568
lubelskie	6452
lubuskie	4798
łódzkie	6145
małopolskie	9197
mazowieckie	9651
opolskie	4903
podkarpackie	5622
podlaskie	6082
pomorskie	8067
śląskie	5005
świętokrzyskie	5880
warmińsko-mazurskie	5363
wielkopolskie	6350
zachodniopomorskie	6950

Rozkład wartości na mapie Polski:



Mapa pozwala zauważyć znaczące różnice cen mieszkań w zależności od regionu, przy czym najdroższe województwa znajdują się w województwach w których znajdują się największe miasta w Polsce. Widać również, że województwa na zachodzie Polski są średnio droższe niż na wschodzie.

Globalne miary autokorelacji przestrzennej

Statystyka *I* Morana

Statystyka *I* Morana jest miarą przestrzennej autokorelacji, która pozwala ocenić, czy wartości danej zmiennej, w tym przypadku ceny za metr kwadratowy, wykazują podobieństwo w przestrzeni.

Wysoka wartość dodatnia statystyki I Morana sugeruje pozytywną autokorelację przestrzenną, co oznacza, że podobne wartości występują blisko siebie (klastry wysokich lub niskich wartości). Wartości ujemne natomiast sugerują negatywną autokorelację, gdzie różne wartości są zlokalizowane w sąsiedztwie siebie. W przypadku braku autokorelacji statystyka powinna być bliska zeru, co wskazuje na losowy rozkład wartości w przestrzeni.

Macierz sąsiedztwa rzędu 1

Macierz sąsiedztwa rzędu 1:

```
$I_observed  
Moran I statistic  
-0.2460021
```

```
$E_I  
Expectation  
-0.06666667
```

```
$Var_I  
Variance  
0.01893216
```

```
$Z_calculated  
Moran I statistic  
-1.303365
```

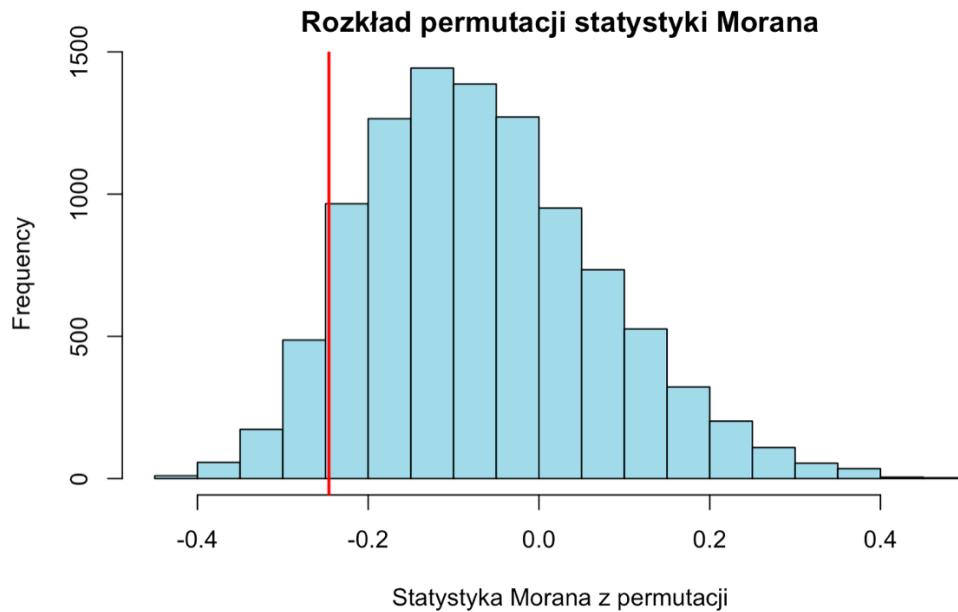
```
$p_value  
[1] 0.9194081
```

Interpretacja:

Statystyka I Morana dla macierzy sąsiedztwa rzędu 1 wynosi -0.246, co sugeruje niewielką, negatywną autokorelację przestrzenną. Oznacza to, że województwa o różnych wartościach ceny za metr kwadratowy nie są skoncentrowane blisko siebie.

Wysokie p-value (0.9194) wskazuje na brak istotności statystycznej, co oznacza, że wynik ten jest najprawdopodobniej przypadkowy.

Statystyka Z wynosząca -1.303 oznacza, że wartość obserwowana statystyki I Morana (-0.246) jest niższa niż wartość oczekiwana, ale różnica ta nie jest statystycznie istotna.



Macierz sąsiedztwa rzędu 4

Macierz sąsiedztwa rzędu 4:

`$I_observed`

Moran I statistic

-0.1200045

`$E_I`

Expectation

-0.06666667

`$Var_I`

Variance

0.00167244

`$Z_calculated`

Moran I statistic

-1.304249

`$p_value`

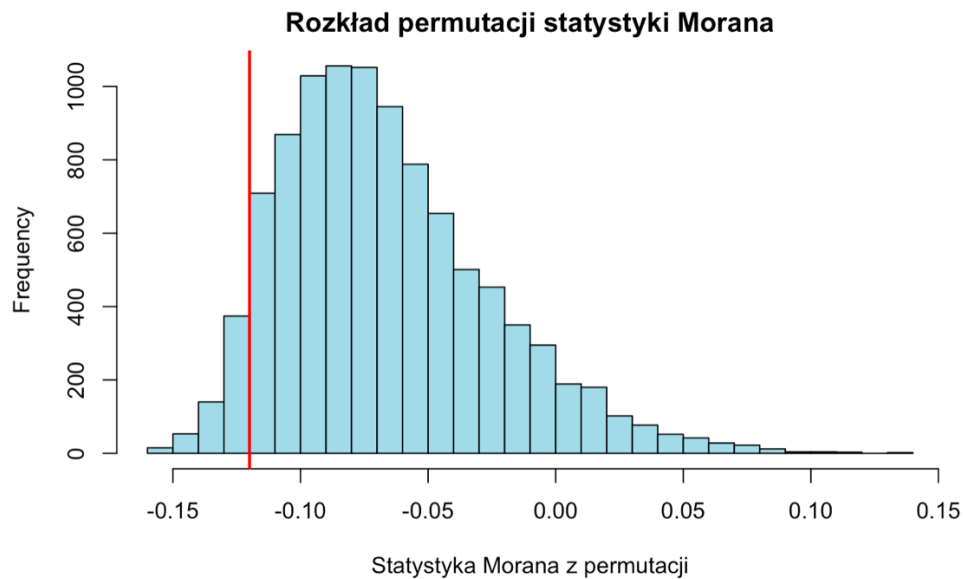
[1] 0.9418058

Interpretacja:

Statystyka I Morana dla macierzy sąsiedztwa rzędu 4 jest mniejsza i wynosi -0.120, co wskazuje na słabą, negatywną autokorelację przestrzenną.

p-value wynoszące 0.9418 również nie jest statystycznie istotne, co oznacza, że nie ma dowodów na autokorelację przestrzenną w przypadku tej macierzy sąsiedztwa.

Statystyka Z wynosząca -1.304 wskazuje, że obserwowana wartość statystyki I Morana (-0.120) jest również niższa od wartości oczekiwanej, ale nie w sposób istotny statystycznie.



Macierz sąsiedztwa oparta na centroidach

Macierz sąsiedztwa oparta na centroidach:

```
$I_observed  
Moran I statistic  
-0.1135192
```

```
$E_I  
Expectation  
-0.06666667
```

```
$Var_I  
Variance  
0.001655381
```

```
$Z_calculated  
Moran I statistic  
-1.151553
```

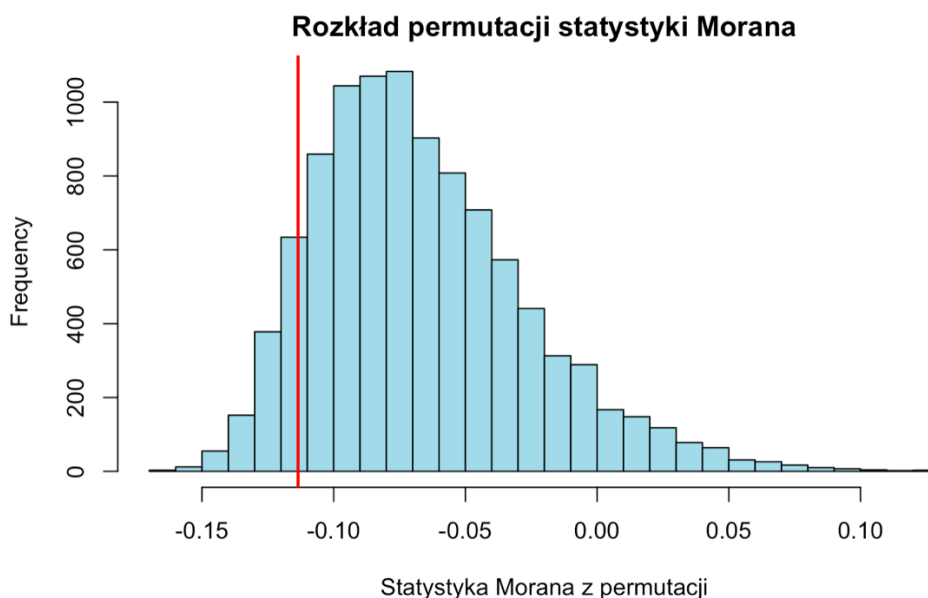
```
$p_value  
[1] 0.9029097
```

Interpretacja:

Dla macierzy opartej na centroidach statystyka Morana wynosi -0.1135, co również wskazuje na niewielką, negatywną autokorelację przestrzenną, podobnie jak w przypadku wcześniejszych macierzy.

p-value wynoszące 0.9029 wskazuje na brak istotności statystycznej, co oznacza, że wynik ten nie jest znaczący i może być przypadkowy.

Wartość Z wynosząca -1.152 oznacza, że wartość obserwowana statystyki I Morana (-0.114) jest nieco niższa od wartości oczekiwanej, jednak ta różnica również nie jest istotna statystycznie.



Podsumowanie

Dla wszystkich analizowanych macierzy sąsiedztwa wyniki wskazują na brak istotnej autokorelacji przestrzennej. Ujemne wartości statystyki Morana sugerują słabą ujemną autokorelację, co może oznaczać, że wartości ceny za metr kwadratowy nie są przestrzennie skoncentrowane w klastrach wysokich lub niskich wartości. Jednakże wysokie p-value we wszystkich przypadkach sugeruje, że te wyniki nie są statystycznie istotne, co wskazuje, że rozkład cen jest losowy i nie wykazuje wyraźnych przestrzennych wzorców.

Statystyka C Geary'ego

Globalna statystyka C Geary'ego mierzy, czy podobne wartości danej zmiennej są skoncentrowane w przestrzeni na poziomie globalnym, obejmując cały analizowany obszar. Analiza ta pozwala ocenić, czy w badanym regionie występuje istotna autokorelacja przestrzenna – wskazuje na ogólne wzorce przestrzenne, gdzie podobne wartości są bardziej skłonne do występowania w bliskim sąsiedztwie, co sugeruje istnienie klastrów wartości podobnych lub odmiennych. W odróżnieniu od lokalnych wersji wskaźników autokorelacji przestrzennej, globalna statystyka C Geary'ego koncentruje się na całym zbiorze danych, nie umożliwiając identyfikacji konkretnych lokalizacji z wysokimi lub niskimi wartościami. Pozwala natomiast stwierdzić, czy dane w ogóle tworzą przestrzennie uporządkowane wzorce, co może być istotne dla analizy trendów i zależności przestrzennych na szeroką skalę.

Opis i interpretacja:

- Test ten bada, czy wartości analizowanej zmiennej (np. średnia cena za m^2) wykazują tendencję do przestrzennej koncentracji, czyli czy skupiają się w jednym lub kilku obszarach. Taka koncentracja może wskazywać na ogólną autokorelację przestrzenną.
- Wysokie dodatnie wartości testu sugerują globalną koncentrację wysokich wartości (globalny „hotspot”), co oznacza, że wyższe wartości są zlokalizowane blisko siebie w skali całego obszaru badania. Z kolei ujemne wartości oznaczają globalną koncentrację niskich wartości (globalny „coldspot”).

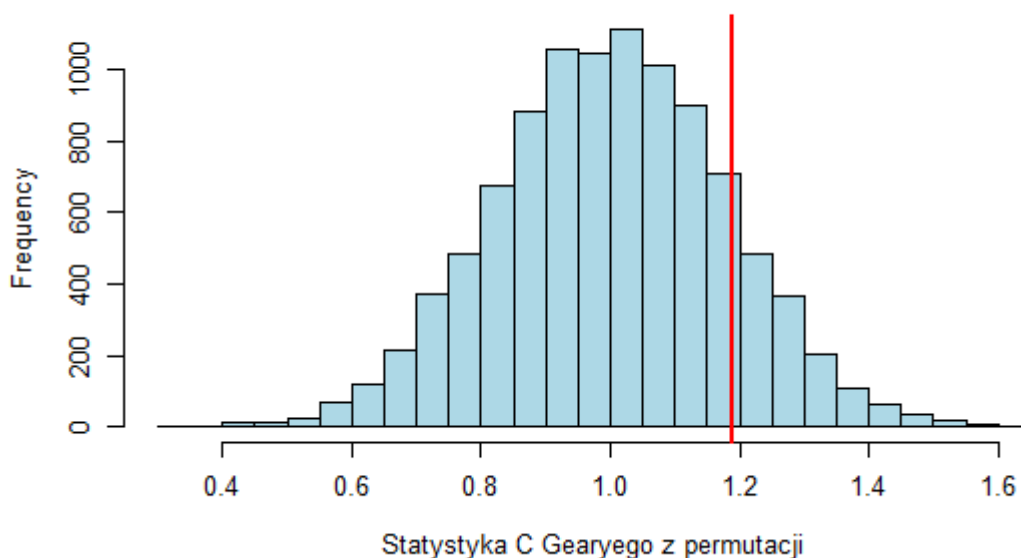
- Wynik testu bliski zeru wskazuje na brak autokorelacji przestrzennej, co oznacza, że wartości są rozmieszczone losowo w przestrzeni i nie tworzą wyraźnych wzorców.

Macierz sąsiedztwa rzędu I

```
"statystyka global c: 1.18796627961646"
"Wartość oczekiwana E[C]: 1"
"Wariancja Var[C]: 0.0316342478531316"
"statystyka Z (obliczona ręcznie): 1.05682040279472"
"statystyka Z (z geary.test): -1.05682040279472"
"p-value dla testu permutacyjnego: 0.142385761423858"
```

Wynik testu globalnego statystyki C Geary'ego wynosi 1.188, co wskazuje na nieznaczny poziom ujemnej autokorelacji przestrzennej w analizowanej zmiennej. Wartość ta sugeruje, że sąsiednie jednostki przestrzenne mają tendencję do różnicowania się pod względem wartości zmiennej, co może oznaczać, że wysokie i niskie wartości są bardziej rozproszone niż skoncentrowane. Jednakże, wynik testu nie jest statystycznie istotny ($p\text{-value} = 0.1424$), co oznacza, że zaobserwowany rozkład może być losowy i nie wskazuje na istotne zjawiska przestrzenne wpływające na rozkład zmiennej „ogółem”.

Rozkład permutacji statystyki Global C Geary'ego



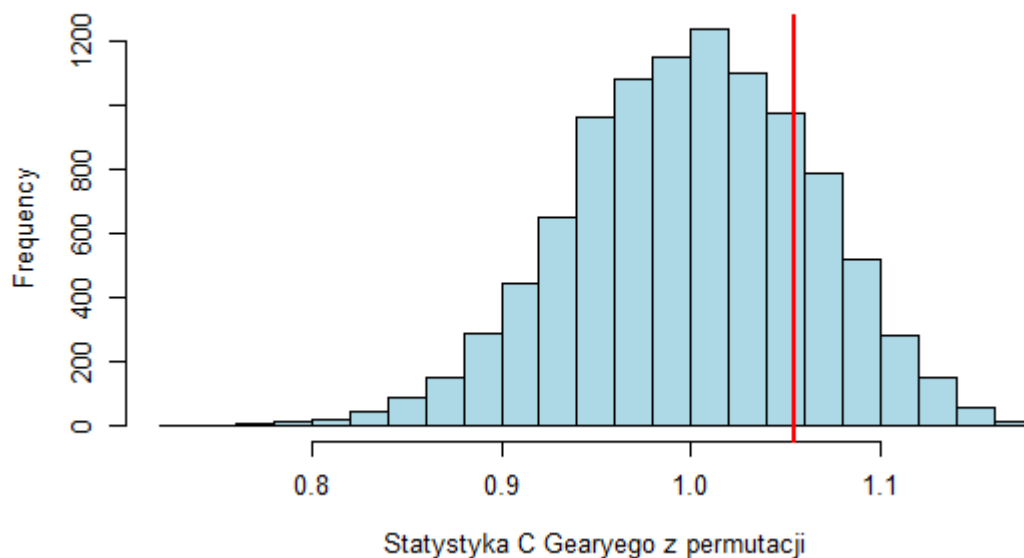
Dla tego testu statystyka Z wynosi -1.0568, co mieści się w przedziale od -1.96 do 1.96. Oznacza to, że wartość statystyki C Geary'ego nie różni się w sposób istotny od wartości oczekiwanych uzyskanych w wyniku losowych permutacji. Statystyka C Geary'ego (ok. 1.188) jest bliska wartości oczekiwanej, co wskazuje, że odchylenie od wartości oczekiwanej jest niewielkie.

Macierz sąsiedztwa rzędu IV

```
"statystyka Global C: 1.05508065781519"  
"wartość oczekiwana E[C]: 1"  
"wariancja var[C]: 0.00402821293866267"  
"statystyka Z (obliczona ręcznie): 0.867846483897053"  
"statystyka Z (z geary.test): -0.867846483897053"  
"p-value dla testu permutacyjnego: 0.203479652034797"
```

Wynik testu globalnej statystyki C Geary'ego wynosi 1.055, co sugeruje brak wyraźnej autokorelacji przestrzennej w analizowanej zmiennej. Wartość ta wskazuje na tendencję do niewielkiej autokorelacji, gdzie sąsiednie jednostki przestrzenne mogą wykazywać tendencję do różnicowania się pod względem wartości zmiennej, co oznacza, że występuje większe rozproszenie niż koncentracja wysokich i niskich wartości. Jednakże, wynik testu nie jest statystycznie istotny ($p\text{-value} = 0.2035$), co sugeruje, że zaobserwowany rozkład zmiennej może wynikać z przypadku, a brak istotnych zjawisk przestrzennych, które wpływają na jej rozmieszczenie.

Rozkład permutacji statystyki Global C Geary'ego

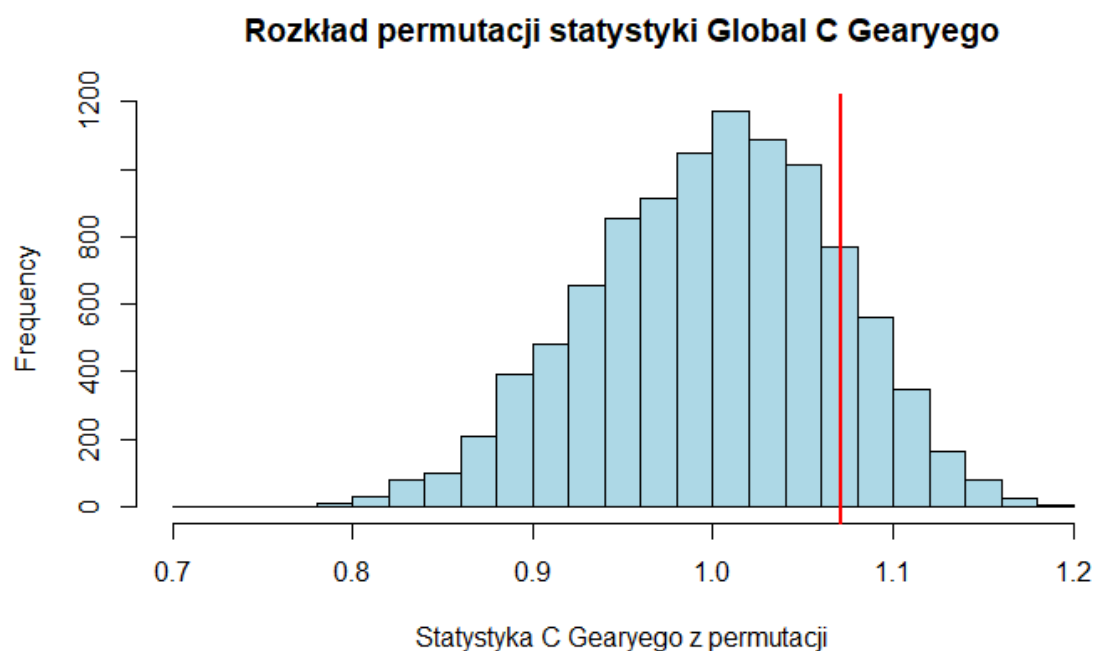


Wnioski po rozkładzie permutacji są bardzo podobne to tych kiedy były obliczane na macierzy wag sąsiedztwa 4 rzędu. Statystyka mieści się pomiędzy -1.96 i 1.96. Oznacza to, że wartość statystyki C Geary'ego nie różni się w sposób istotny od wartości oczekiwanych uzyskanych w wyniku losowych permutacji.

Macierz sąsiedztwa oparta na centroidach

```
"Statystyka Global C: 1.07130902368427"  
"wartość oczekiwana E[C]: 1"  
"wariancja var[C]: 0.00462290157629001"  
"Statystyka Z (obliczona ręcznie): 1.0487866891922"  
"Statystyka Z (z geary.test): -1.0487866891922"  
"p-value dla testu permutacyjnego: 0.147985201479852"
```

Wynik testu globalnej statystyki C Geary'ego wynosi 1.071, co sugeruje brak wyraźnej autokorelacji przestrzennej w analizowanej zmiennej. Wartość ta wskazuje na niewielką tendencję do autokorelacji, co oznacza, że sąsiednie jednostki przestrzenne mogą wykazywać pewne różnice w wartościach zmiennej, lecz te różnice są stosunkowo małe. Zjawisko to może sugerować, że wartości zmiennej są rozproszone, ale nie wykazują silnych skupisk ani klastrów wysokich lub niskich wartości. Niemniej jednak, wynik testu nie jest statystycznie istotny ($p\text{-value} = 0.148$), co wskazuje na to, że zaobserwowany rozkład zmiennej może być przypadkowy, a brak istotnych efektów przestrzennych nie pozwala na wyciąganie wniosków o wpływie przestrzeni na jej rozmieszczenie.



Wykres jest niemal identyczny do poprzedniego, stąd wnioski można wyciągnąć dokładnie te same.

Statystyka G Getisa i Orda

Globalna statystyka Getisa i Orda mierzy, czy wysokie lub niskie wartości danej zmiennej są skoncentrowane w przestrzeni na poziomie globalnym (dla całego obszaru badania). Analiza ta pozwala ocenić, czy w całym badanym obszarze występuje istotna autokorelacja przestrzenna dla wartości wysokich lub niskich, wskazując na ogólne trendy przestrzenne. W odróżnieniu od wersji lokalnej wskaźnika, wersja globalna skupia się na całym zbiorze danych i nie pozwala na lokalizowanie specyficznych obszarów o niskich lub wysokich wartościach, ale wskazuje, czy w ogóle występują takie wzorce.

Opis i interpretacja:

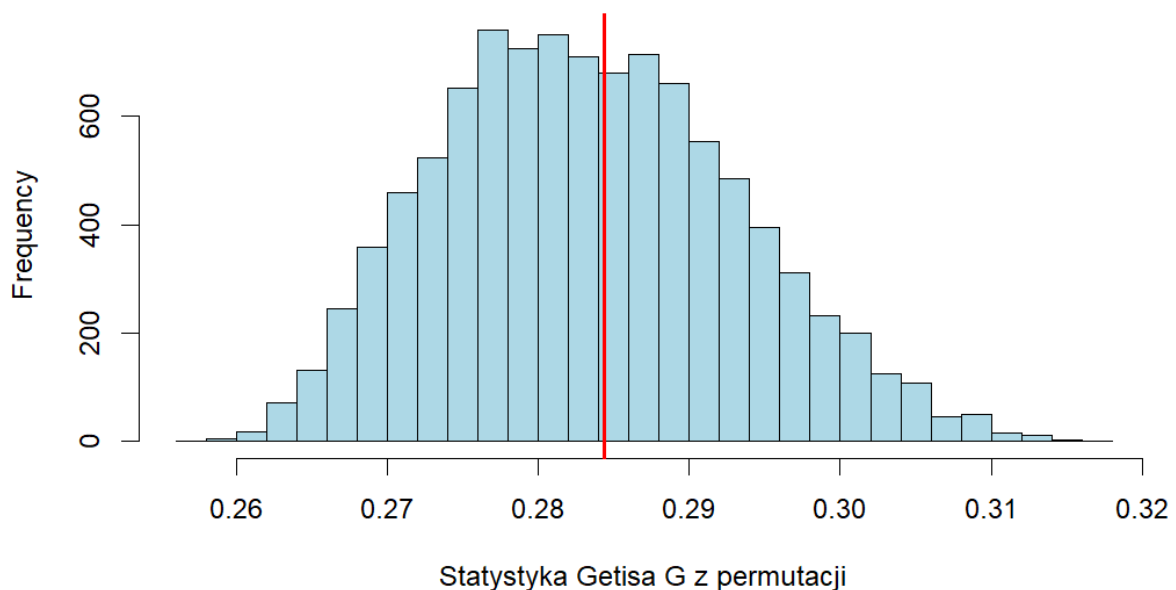
- Test ten analizuje, czy wartości zmiennej (np. średnia cena za m²) są przestrzennie skoncentrowane w jednym lub kilku obszarach, co może sugerować ogólną tendencję do autokorelacji przestrzennej.
- Wysokie wartości dodatnie testu wskazują na globalną koncentrację wysokich wartości (czyli ogólny „hotspot”), podczas gdy wartości ujemne oznaczają globalną koncentrację niskich wartości (czyli globalny „coldspot”).
- Wartość testu bliska zeru sugeruje brak globalnej autokorelacji, czyli przypadkowy rozkład wartości w przestrzeni.

Macierz sąsiedztwa rzędu 1

```
"Statystyka Global G: 0.284371429680849"  
"wartość oczekiwana E[G]: 0.283333333333333"  
"Wariancja Var[G]: 9.54366915199845e-05"  
"Statystyka Z (obliczona ręcznie): 0.106262486823169"  
"Statystyka Z (z globalG.test): 0.106262486823169"  
"p-value dla testu permutacyjnego: 0.445655434456554"
```

Wynik testu globalnego statystyki G wynosi 0.2844, wskazuje to na umiarkowany poziom przestrzennej autokorelacji w analizowanej zmiennej. Wartość ta sugeruje, że istnieje pewna tendencja do grupowania się wysokich (lub niskich) wartości w przestrzeni, co może świadczyć o zjawiskach lokalnych lub regionalnych wpływających na rozkład zmiennej „ogółem”.

Rozkład permutacji statystyki Global G Getisa



Wynik testu permutacyjnego dla statystyki G Getisa i Orda, z wartością p równą 0.4457, sugeruje, że nie ma istotnej autokorelacji przestrzennej w analizowanej zmiennej.

Statystyka Z w testowaniu hipotez:

$Z > 1.96$ oznacza, że wynik jest istotny statystycznie, tzn. istnieje silne odchylenie od wartości oczekiwanej.

$Z < -1.96$ również oznacza istotne odchylenie, ale w przeciwną stronę.

Z pomiędzy -1.96 a 1.96 sugeruje, że wynik nie jest istotny statystycznie, czyli brak wystarczających dowodów na odchylenie od oczekiwanej wartości.

Wartość Z wynosząca **0.106** oznacza, że statystyka G jest bardzo bliska wartości oczekiwanej, ponieważ Z wskazuje, jak daleko (w jednostkach odchylenia standardowego) znajduje się G od $E[G]$. Ponieważ wartość Z jest bardzo mała (bliska zeru), oznacza to, że G nie odbiega znacząco od wartości oczekiwanej, co sugeruje, że dane są w miarę jednorodne lub nie wykazują silnej koncentracji przestrzennej.

Oznacza to, że obserwowana statystyka G nie różni się w istotny sposób od wartości uzyskanych w wyniku permutacji, co sugeruje, że rozkład wartości „ogółem” jest losowy, bez wyraźnych skupień w przestrzeni. Wyniki są nieistotne statystycznie, więc nie ma dowodów na istnienie istotnej przestrzennej koncentracji w danych dotyczących średniej ceny mieszkań za metr kwadratowy.

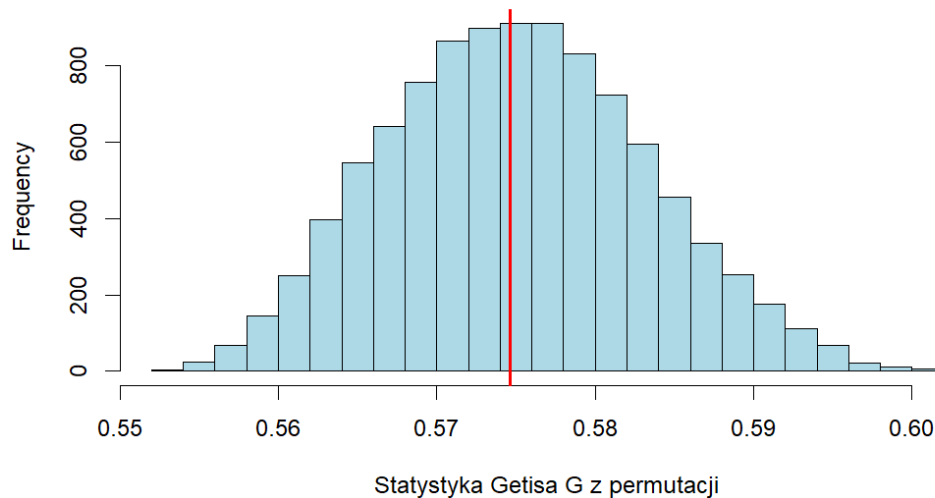
Macierz sąsiedztwa rzędu IV

```
[1] "Statystyka global G: 0.574611298161974"
[1] "Wartość oczekiwana E[G]: 0.575"
[1] "Wariancja Var[G]: 6.59043122504532e-05"
[1] "Statystyka Z (obliczona ręcznie): -0.0478806103774334"
[1] "Statystyka Z (z globalG.test): -0.0478806103774334"
[1] "p-value dla testu permutacyjnego: 0.511348865113489"
```

Wynik testu globalnej statystyki G dla zmiennej „ogółem” z użyciem macierzy sąsiedztwa rzędu 4 wskazuje na wartość statystyki G równą 0.5746. Wartość standardowego odchylenia wynosi -0.0479. Wartość p wynosząca 0.5191 jest znacznie wyższa niż typowy próg istotności (0.05), co oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej.

To wskazuje na brak istotnej autokorelacji przestrzennej w analizowanej zmiennej w kontekście sąsiedztwa rzędu 4.

Rozkład permutacji statystyki Global G Getisa



Wynik testu permutacyjnego dla statystyki G Getisa i Orda, z wartością p równą 0.5113, wskazuje na brak istotnej autokorelacji przestrzennej w analizowanej zmiennej.

Wynik statystyki Z równy -0.047 sugeruje, że nie ma wystarczających dowodów, by odrzucić hipotezę zerową, co oznacza, że brak jest przestrzennych zależności w analizowanych danych.

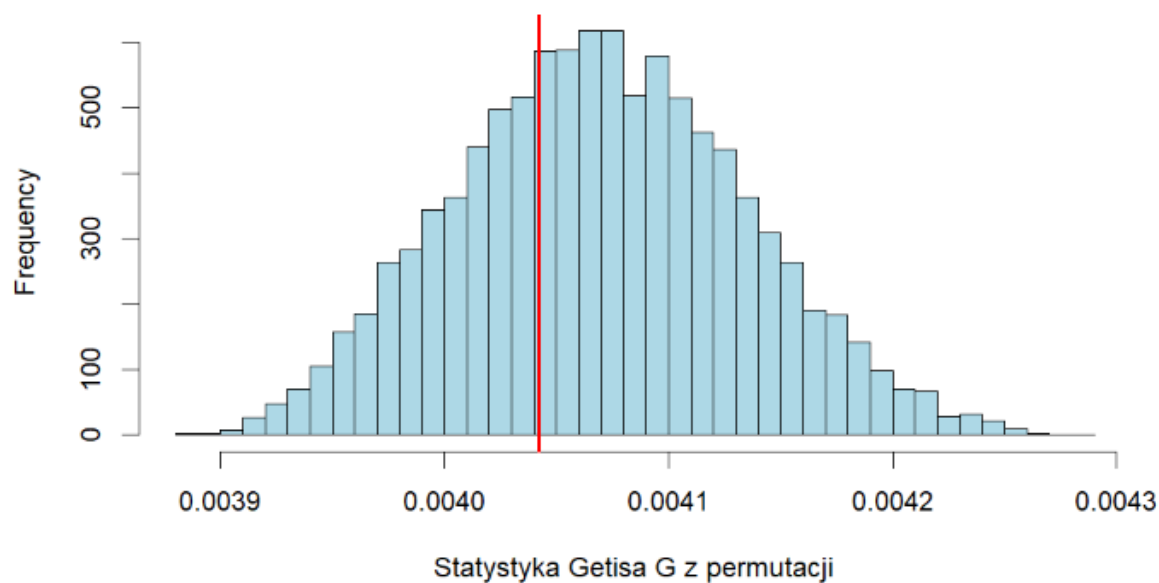
Macierz sąsiedztwa oparta na centroidach

```
[1] "Statystyka Global G: 0.00404194417153501"
[1] "wartość oczekiwana E[G]: 0.00406939869868106"
[1] "wariancja var[G]: 4.07215661025155e-09"
[1] "Statystyka Z (obliczona ręcznie): -0.430231031864512"
[1] "Statystyka Z (z globalG.test): -0.430231031864512"
[1] "p-value: 0.658834116588341"
```

Wynik testu globalnej statystyki G, bazujący na macierzy sąsiedztwa opartej na centroidach dla zmiennej „ogółem”, wskazuje na wartość statystyki G równą 0.0040. Wartość p wynosząca 0.6665 jest znacznie wyższa od typowego poziomu istotności 0.05, co oznacza, że nie możemy odrzucić hipotezy zerowej.

W związku z tym, brak istotnej autokorelacji przestrzennej w analizowanej zmiennej sugeruje, że rozkład wartości „ogółem” nie wykazuje wyraźnych skupień w przestrzeni. Może to wskazywać na losowy charakter rozkładu tej zmiennej, bez widocznych wzorców przestrzennych.

Rozkład permutacji statystyki Global G Getisa



Wynik testu permutacyjnego dla statystyki G Getisa i Orda, z wartością p równą 0.6588, wskazuje na brak istotnej autokorelacji przestrzennej w analizowanej zmiennej.

Podobnie jak w poprzednich przypadkach wynik statystyki Z równy -0.043 sugeruje, że nie ma wystarczających dowodów, by odrzucić hipotezę zerową, co oznacza, że brak jest przestrzennych zależności w analizowanych danych.

Korelogram i wykres rozproszenia

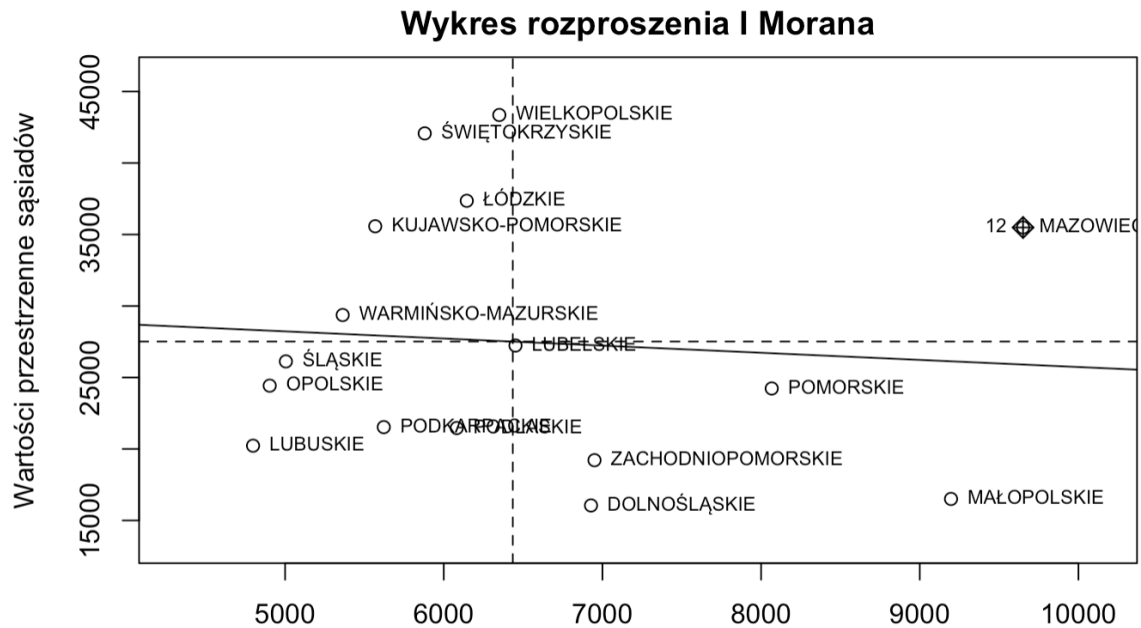


Na Korelogramie I Morana przedstawiono wartość statystyki I Morana w funkcji różnych opóźnień (lags). Oś pozioma reprezentuje kolejne opóźnienia (lags), natomiast oś pionowa pokazuje wartość statystyki I Morana. Punkty i odpowiadające im słupki błędów wskazują na autokorelację przestrzenną dla różnych poziomów sąsiedztwa.

Interpretacja:

Wartości I Morana bliskie zera lub ujemne sugerują brak autokorelacji przestrzennej lub ujemną korelację, co oznacza, że ceny za metr kwadratowy mogą być zróżnicowane pomiędzy sąsiednimi województwami.

Pierwsze opóźnienie wykazuje ujemną korelację, co może sugerować, że województwa sąsiadujące mają różniące się ceny. Dalsze opóźnienia mają wartości bliskie zeru, co wskazuje na zanikanie korelacji przestrzennej z dalszym sąsiedztwem.



Wykres rozproszenia I Morana

Na wykresie rozproszenia I Morana widzimy analizę przestrzenną dla średnich cen za metr kwadratowy w województwach w Polsce.

Oś pozioma przedstawia średnie ceny za metr kwadratowy w poszczególnych województwach, natomiast oś pionowa przedstawia przestrzenną średnią cen dla sąsiednich województw.

Punkty reprezentują województwa, a ich rozmieszczenie względem osi krzyżujących się na środku wykresu wskazuje na relację między wartościami ceny a przestrzennymi średnimi.

Etykiety województw umożliwiają identyfikację, które z nich mają wyższe lub niższe wartości cen oraz czy ich ceny są podobne do cen sąsiednich województw.

Interpretacja: Województwa z wysokimi cenami za metr kwadratowy, takie jak Mazowieckie, znajdują się w prawym górnym rogu, co może sugerować wyższe wartości również w sąsiednich województwach.

Z kolei województwa w lewym dolnym rogu (takie jak Lubuskie) mają niższe ceny oraz są otoczone innymi województwami o stosunkowo niższych cenach.

Linia regresji wskazuje ogólny trend zależności, a jej nachylenie pokazuje czy istnieje pozytywna lub negatywna autokorelacja przestrzenna.

Wspólna Interpretacja

Analiza przestrzennej autokorelacji średnich cen mieszkań za metr kwadratowy w Polsce nie wykazała istotnych wzorców przestrzennych. Globalne miary autokorelacji, takie jak statystyka I Morana, C Geary'ego, oraz G Getisa i Orda, sugerują brak istotnych klastrów wysokich lub niskich

wartości cenowych w sąsiednich województwach. Wyniki są zgodne dla różnych typów macierzy sąsiedztwa (rzędu I, IV oraz opartych na centroidach), a wartości statystyk nie są statystycznie istotne, co potwierdzają wysokie wartości p oraz małe wartości statystyk Z .

Korelogram I Morana pokazuje brak istotnej autokorelacji przestrzennej przy różnych opóźnieniach, co sugeruje, że ceny mieszkań w sąsiednich województwach nie są ze sobą silnie powiązane. Wykres rozproszenia I Morana również wskazuje, że województwa z wysokimi cenami (np. Mazowieckie) sąsiadują z województwami o różnorodnych wartościach cenowych, a niektóre województwa o niższych cenach (np. Lubuskie) są otoczone regionami o podobnie niskich cenach.

Podsumowanie

Brak wyraźnej przestrzennej autokorelacji wskazuje na to, że ceny mieszkań w Polsce mogą być rozproszone losowo, bez wyraźnych klastrów regionalnych. W związku z tym, wpływ przestrzennej zależności na średnie ceny za metr kwadratowy jest znikomy, co może sugerować, że inne czynniki (np. lokalna gospodarka, infrastruktura) mają silniejszy wpływ na ceny niż sama bliskość geograficzna.

Podsumowując analizę, można zauważyć, że rozkład cen mieszkań w Polsce różni się w zależności od regionu. Najdroższe mieszkania znajdują się w województwach z dużymi ośrodkami miejskimi, takimi jak Mazowieckie (Warszawa) oraz Małopolskie (Kraków). To w tych regionach ceny za metr kwadratowy są najwyższe, co wynika z koncentracji miejsc pracy, uczelni oraz rozwiniętej infrastruktury.

Z kolei najniższe ceny mieszkań występują w województwach takich jak Lubuskie czy Podkarpackie, gdzie presja cenowa jest mniejsza ze względu na mniejsze aglomeracje i słabszy rozwój gospodarczy. W tych regionach dostęp do niektórych usług i infrastruktury może być ograniczony, co może wpływać na mniejsze zapotrzebowanie na mieszkania i, w konsekwencji, niższe ceny.