



AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

Wydział Zarządzania

Ekonometria Przestrzenna

Projekt 1

Jakub Le Van, Mateusz Mulka

Kraków 31.10.2024 r.

Spis treści

1. Wstęp	3
2. Autokorelacja przestrzenna	3
3. Autokorelacja z permutacjami	5
4. Wykres rozproszenia i korelogram.....	6
5. Podsumowanie	8

1. Wstęp

W projekcie zajmiemy się analizą autokorelacji przestrzennej danych dotyczących odpadów komunalnych zebranych w roku 2022 na poziomie województw. Przestrzenna autokorelacja odzwierciedla, w jakim stopniu wartości danej zmiennej w poszczególnych lokalizacjach są powiązane z wartościami tej zmiennej w sąsiednich obszarach. W kontekście odpadów komunalnych, analiza ta pozwoli nam zidentyfikować potencjalne wzorce i zależności przestrzenne między województwami w zakresie ilości zbieranych odpadów, co może dostarczyć istotnych informacji dla polityki zarządzania gospodarką odpadami.

2. Autokorelacja przestrzenna

W analizie autokorelacji przestrzennej badamy, czy i w jakim stopniu wartości badanej zmiennej – w tym przypadku ilości odpadów komunalnych zebranych w województwach w roku 2022 – wykazują przestrzenną zależność. W użytych testach mamy następujące hipotezy:

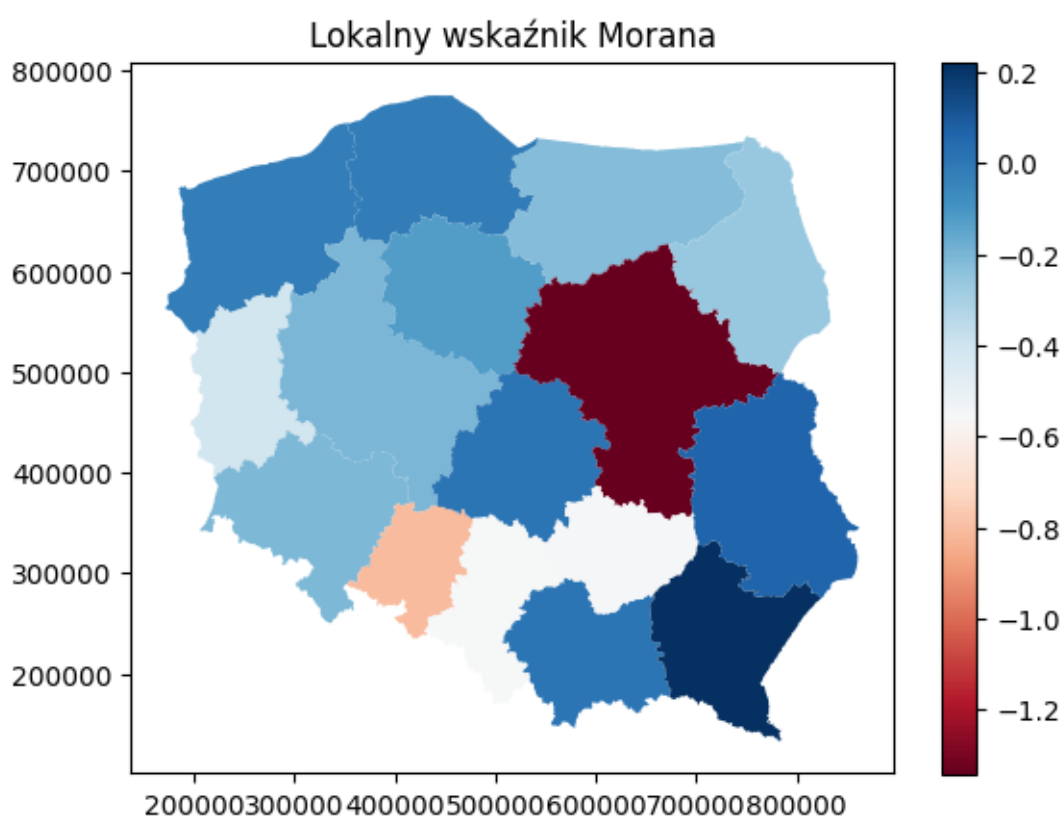
- H_0 – Brak autokorelacji przestrzennej
- H_1 – Istnieje autokorelacja przestrzenna

W celu zbadania przestrzennej struktury zależności przeanalizowaliśmy wskaźnik Morana I, wskaźnik Geary'ego C oraz statystykę G Getisa i Orda, które wyglądają następująco:

	WSKAŹNIK	STATYSTYKA Z	P-VALUE
MORAN I	-0.29494410674958277	-1.5346937902035942	0.12485906191611987
GEARY C	1.2736093887115443	1.8208863903188468	0.03431206420540975
G GETIS I ORD	0.29029858981786816	0.23627946059469418	0.4066079119212024

Wskaźnik Morana I wykazał umiarkowanie negatywną autokorelację, co sugeruje brak istotnych podobieństw między sąsiadującymi województwami w zakresie badanej zmiennej. Jednak z uwagi na wartość p-value większą od 0.05 nie mamy podstaw na odrzucenie hipotezy zerowej. Na podstawie tego wskaźnika musimy zatem przyjąć brak autokorelacji przestrzennej. Warto dodać tutaj, że wskaźnik Morana I mierzy autokorelację globalną uwzględniającą ogólny wzorec w danych przestrzennych, podczas gdy pozostałe wskaźniki mają bardziej charakter

lokalny. Wartości wskaźnika Geary'ego C również wskazują na zróżnicowanie przestrzenne (negatywna autokorelacja przestrzenna), tym razem p-value wynosi 0.03, a więc mamy podstawy do takiego stwierdzenia. W przypadku statystyki G Getisa i Orda, z uwagi na wysoką wartość p-value nie możemy odrzucić hipotezy zerowej. Oznacza to, że nie wykryto statystycznie istotnych skupisk wysokich lub niskich wartości w danych, co sugeruje, że zmienna jest rozłożona losowo i brak jest przestrzennych wzorców w rozkładzie analizowanej zmiennej.



W przypadku lokalnego wskaźnika morana ciemniejsze odcienie czerwieni wskazują na silną negatywną autokorelację, co oznacza, że województwa te znacząco różnią się od sąsiadów pod względem wartości zmiennej (np. województwo lubelskie i podkarpackie). Jaśniejsze odcienie niebieskiego, widoczne głównie na północy, sugerują brak wyraźnych wzorców autokorelacji, co może oznaczać większą jednorodność lub niezależność od sąsiednich województw. Mapa wskazuje na lokalne zróżnicowanie w gospodarce odpadami.

3. Autokorelacja z permutacjami

W tej części przeanalizujemy istotność autokorelacji przestrzennej przy użyciu testów permutacyjnych dla wskaźników Morana I, Geary'ego C oraz statystyki G Getisa i Orda.

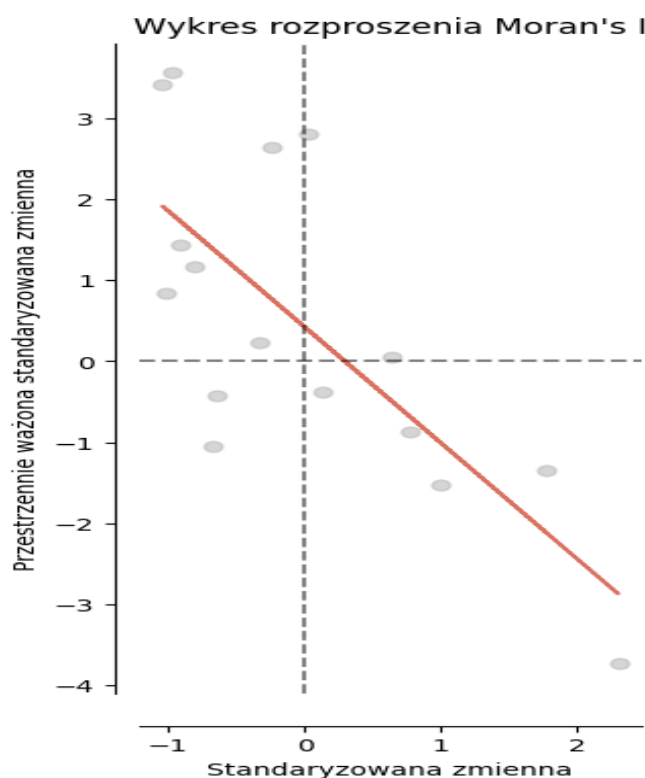
	WSKAŹNIK	STATYSTYKA Z (Z PERMUTACJĄ)	P-VALUE (Z PERMUTACJĄ)
MORAN I	-0.29494410674958277	-1.5531629492238848	0.043
GEARY C	1.2736093887115443	-	0.025
G GETIS I ORD	0.29029858981786816	-	0.412

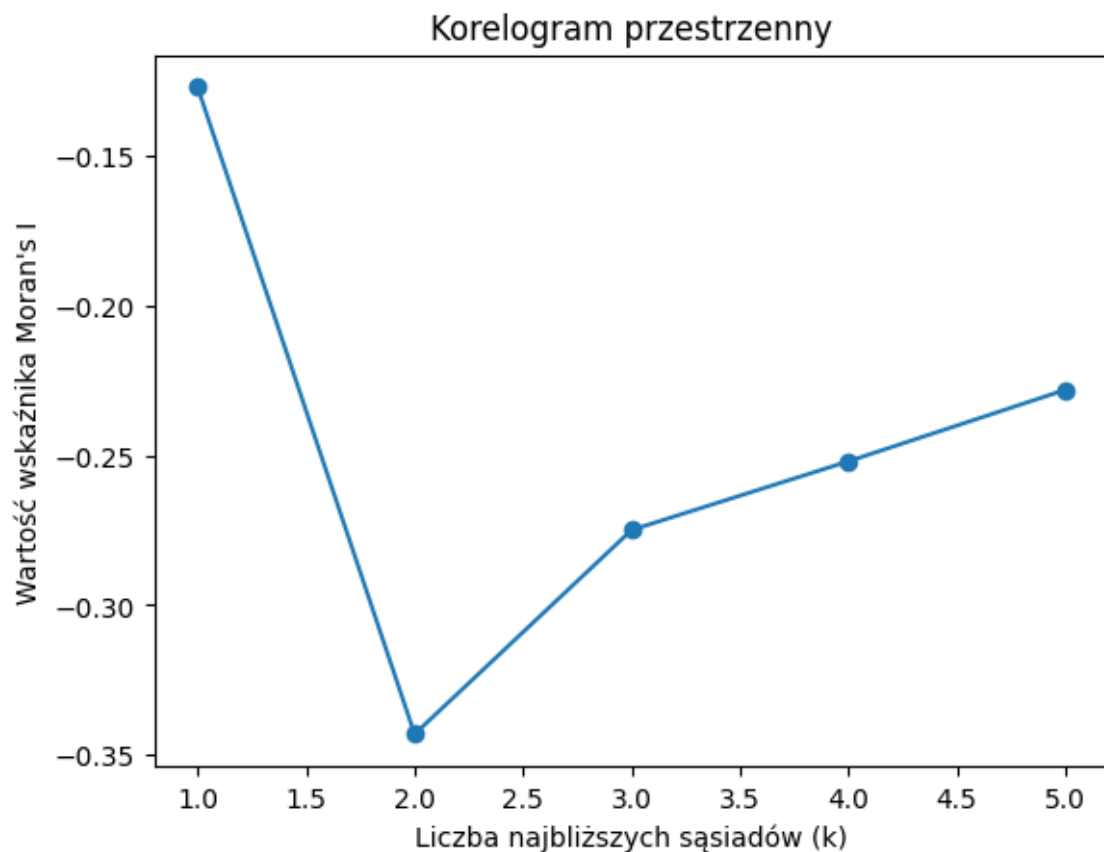
Wskaźnik Morana I, z wartością $p=0,043$, wskazuje na istotną statystycznie negatywną autokorelację przestrzenną, co sugeruje pewne zróżnicowanie wartości zmiennej między sąsiednimi województwami. Wskaźnik Geary'ego C, z wartością $p=0,025$, potwierdza negatywną autokorelację, wskazując na wyraźne różnice między sąsiednimi regionami. Natomiast statystyka G Getisa i Orda, z wartością $p=0,412$, nie wykazuje istotnych skupisk wysokich ani niskich wartości, co sugeruje losowy charakter rozmieszczenia wartości w danych.

Podsumowując, testy permutacyjne wzmacniają interpretację wyników, potwierdzając umiarkowaną, lecz istotną statystycznie negatywną autokorelację przestrzenną dla wskaźników Morana I i Geary'ego C, podczas gdy statystyka G Getisa i Orda pozostaje nieistotna. Finalnie wskazuje to na obecność zróżnicowania przestrzennego między sąsiednimi województwami w zakresie odpadów komunalnych, przy braku wyraźnych skupisk wartości w skali regionalnej.

4. Wykres rozproszenia i korelogram

Wykres rozproszenia Morana przedstawia zależność między standaryzowaną zmienną a przestrzennie ważoną standaryzowaną zmienną. Na podstawie tego wykresu możemy wysnuć kilka wniosków. Czerwona linia to linia regresji, odpowiada ona wartości wskaźnika Morana. Ponieważ ma ona ujemne nachylenie, wartość wskaźnika Morana jest ujemna, co sugeruje negatywną autokorelację przestrzenną. Oznacza to, że wysokie wartości zmiennej mają tendencję do sąsiadowania z niskimi wartościami, a niskie z wysokimi. Wykres jest też podzielony na cztery ćwiartki. Możemy zauważyć zdecydowane skoncentrowanie punktów w lewej górnej oraz prawej dolnej ćwiartce, są to ćwiartki odpowiadające parom „niskie-wysokie” oraz „wysokie-niskie”, co potwierdza wcześniej wysnute wnioski. Dodatkowo warto zauważyć, że praktycznie brak jest punktów w prawej górnej ćwiartce odpowiadającej za sąsiedztwo wysokich wartości z wysokimi. Możemy zatem zauważyć, że województwa z dużą liczbą odpadów sąsiadują z województwami o niskim natężeniu tego wskaźnika. Jak mogliśmy zaobserwować we wcześniejszej części sprawozdania, najwyższa wartość tego wskaźnika jest dla woj. Mazowieckiego, a więc prawdopodobnie to Warszawa ma duży wpływ na wyniki.





Korelogram przestrzenny dla wskaźnika Morana I wskazuje, że przy jednym najbliższym sąsiedzie wartość autokorelacji jest bliska zero, co oznacza niewielkie podobieństwo lub różnice między sąsiadującymi województwami. Najbardziej ujemna autokorelacja (około -0.35) występuje jednak przy dwóch sąsiadach, co sugeruje, że na tym dystansie województwa wykazują największe zróżnicowanie w zakresie badanej zmiennej (gospodarka odpadami). W miarę dalszego zwiększania liczby sąsiadów (lagów) autokorelacja stopniowo wzrasta, zbliżając się do zera, co oznacza zmniejszające się zróżnicowanie między regionami na większych odległościach. Wynik ten wskazuje na dominację lokalnych czynników w zarządzaniu odpadami, przy czym różnice są najbardziej widoczne na poziomie bliskich sąsiadów.

5. Podsumowanie

W raporcie przeanalizowano przestrzenną autokorelację w danych dotyczących odpadów komunalnych zebranych w 2022 roku na poziomie województw, stosując zarówno klasyczne testy autokorelacji, jak i testy permutacyjne. Wskaźnik Morana I nie wykazał istotnej globalnej autokorelacji przestrzennej, natomiast wskaźnik Geary'ego C zasugerował lokalne zróżnicowanie między województwami, wskazując na negatywną autokorelację przestrzenną. Testy permutacyjne potwierdziły umiarkowaną, statystycznie istotną negatywną autokorelację zarówno dla wskaźnika Morana I, jak i Geary'ego C, przy braku istotnych wyników dla statystyki G Getisa-Orda, co sugeruje losowy rozkład wartości bez znaczących klastrów. Wykres rozproszenia Morana oraz korelogram wskazały na negatywną zależność między wartościami zmiennej w sąsiadujących województwach, zwłaszcza na krótkich dystansach, co podkreśla lokalne zróżnicowanie. Wyniki te sugerują, że gospodarka odpadami w Polsce jest silnie uwarunkowana lokalnymi czynnikami, co może stanowić istotną wskazówkę dla planowania regionalnych polityk zarządzania odpadami.