



AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

Wydział Zarządzania

Ekonometria Przestrzenna

Projekt 2 - Powiaty

Jakub Le Van, Mateusz Mulka

Spis treści

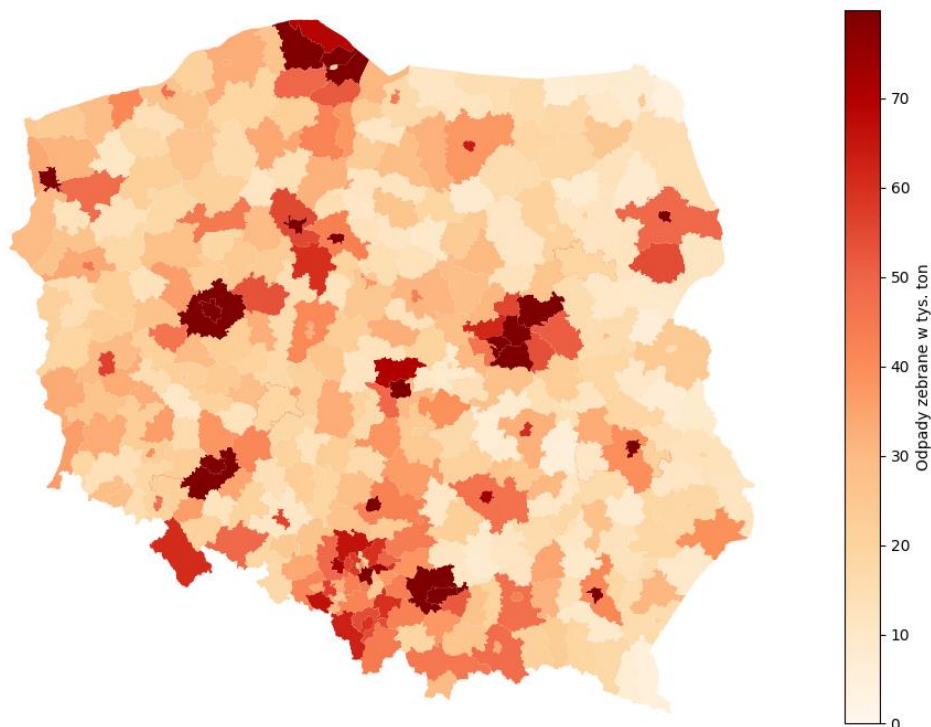
1.	Wstęp	3
2.	Rozkład surowych danych – odpady zebrane w ciągu roku	3
3.	Macierze wag przestrzennych	4
4.	Autokorelacja przestrzenna	5
5.	Testy permutacyjne	7
6.	Wykres rozproszenia i korelogram	8
7.	Analiza wyników dla różnych macierzy odległości	10
8.	Analiza dwóch wybranych województw.....	11
8.1.	Województwo Zachodniopomorskie	11
8.1.1.	Analiza autokorelacji przestrzennej.....	13
8.1.2.	Wykres rozproszenia i korelogram	15
8.1.3	Analiza wyników dla różnych macierzy odległości.....	17
8.2.	Województwo Warmińsko-Mazurskie	18
8.2.1.	Analiza autokorelacji przestrzennej	19
8.2.2.	Wykres rozproszenia i korelogram	22
8.2.3	Analiza wyników dla różnych macierzy odległości	23
9.	Porównanie uzyskanych wyników	24
10.	Porównanie wskaźników autokorelacji dla powiatów i województw (wcześniejszy projekt)	26
11.	Podsumowanie.....	27

1. Wstęp

Projekt dotyczy analizy przestrzennej autokorelacji danych dotyczących odpadów komunalnych zebranych w 2022 roku na poziomie powiatów. Przestrzenna autokorelacja pozwala zbadać, w jakim stopniu wartości danej zmiennej w jednym obszarze są zależne od wartości w otaczających go lokalizacjach. Analizując dane o odpadach komunalnych, będziemy mogli zidentyfikować wzorce oraz zależności przestrzenne między powiatami w zakresie ilości zbieranych odpadów. Wyniki analizy mogą dostarczyć kluczowych informacji, które będą pomocne w optymalizacji zarządzania gospodarką odpadami oraz w podejmowaniu decyzji na poziomie lokalnym i regionalnym.

2. Rozkład surowych danych – odpady zebrane w ciągu roku

Mapa przedstawia przestrzenny rozkład ilości odpadów komunalnych zebranych w 2022 roku na poziomie powiatów w Polsce (w tys. ton). Ciemniejsze odcienie czerwieni wskazują na powiaty o wyższej ilości zebranych odpadów, natomiast jaśniejsze reprezentują niższe wartości.



Najwyższe wartości obserwowane są w powiatach obejmujących duże miasta, takie jak Warszawa, Kraków czy Wrocław, co jest zgodne z większą gęstością zaludnienia. Powiaty wiejskie charakteryzują się niższymi poziomami zebranych odpadów.

3. Macierze wag przestrzennych

W analizie przestrzennej istotnym elementem jest określenie relacji między jednostkami geograficznymi poprzez macierze wag. W projekcie zdefiniowano kilka typów macierzy wag, które odzwierciedlają różne podejścia do modelowania zależności przestrzennych. W głównej części analizy posługujemy się macierzą typu Queen, jednak później obliczamy statystyki dla wszystkich macierzy i porównujemy ze sobą wyniki. Wybór tej macierzy był uzasadniony większą elastycznością w identyfikowaniu przestrzennych zależności, szczególnie w przypadku regionów o nieregularnych kształtach powiatów i nieregularnych odległościach od centroidów.

1) **Macierz sąsiedztwa typu Queen:**

- a) Dwa powiaty są sąsiadami, jeśli dzielą wspólną granicę lub stykają się na wierzchołku.
- b) Macierz typu Queen jest bardziej "globalna" i uwzględnia większą liczbę relacji przestrzennych między jednostkami.

2) **Macierz sąsiedztwa typu Rook:**

- a) Sąsiedztwo w tej macierzy definiowane jest wyłącznie przez wspólne granice między powiatami, bez uwzględniania punktów styku w wierzchołkach.
- b) Podejście to lepiej oddaje "liniowe" zależności przestrzenne.
- c) Warty zauważenia jest fakt, że z uwagi na brak kwadratowych powiatów, możemy spodziewać się takich samych lub bardzo zbliżonych wyników, co w przypadku poprzedniej macierzy odległości.

3) **Macierz odległościowa z progiem:**

- a) Relacje przestrzenne wyrażone są za pomocą odległości między centrami obszarów (centroidami). Wagi maleją wraz z większym dystansem, co odzwierciedla słabnięcie wpływu przestrzennego na większe odległości.
- b) Macierz została skonstruowana z progiem odległości 50 km.

4) Macierz KNN

- a) Macierz KNN definiuje sąsiadów na podstawie odległości od centroidów podobnie jak poprzednio opisana macierz, jednak w tym przypadku użyta jest stała liczba sąsiadów. W przypadku naszego kodu liczba ta została ustalona jako 4.

4. Autokorelacja przestrzenna

W analizie autokorelacji przestrzennej sprawdzamy, czy i w jakim stopniu wartości badanej zmiennej – w tym przypadku ilości odpadów komunalnych zebranych na poziomie powiatów w roku 2022 – wykazują zależności przestrzenne. Analiza ta opiera się na weryfikacji następujących hipotez statystycznych:

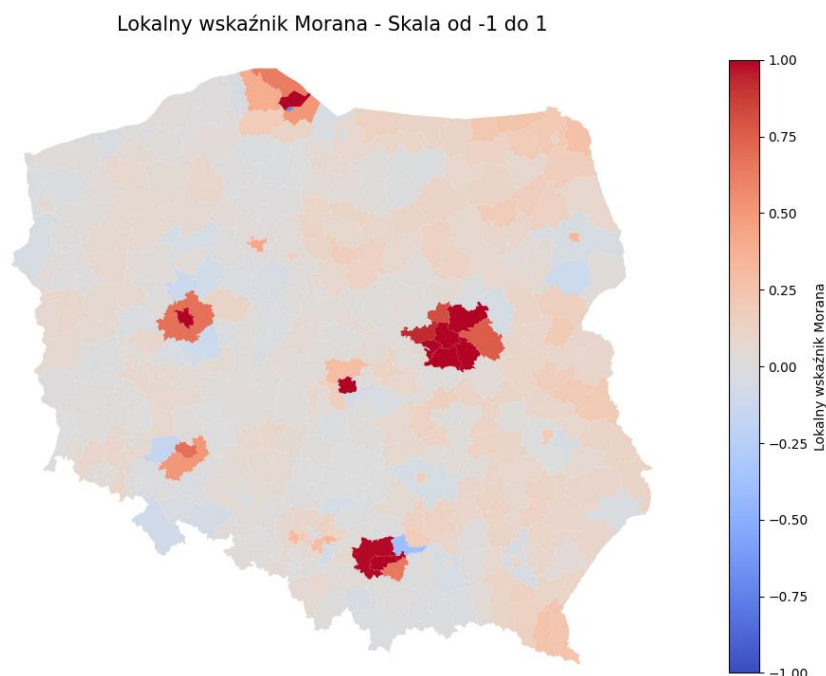
- **H0:** Brak autokorelacji przestrzennej.
- **H1:** Istnieje autokorelacja przestrzenna.

Aby zbadać przestrzenną strukturę zależności, obliczyliśmy wskaźnik Morana I, wskaźnik Geary'ego C oraz statystykę G Getisa i Orda. Wyniki tych miar umożliwiają ocenę, czy obserwowane wartości zmiennej w powiatach są powiązane z wartościami w sąsiednich jednostkach przestrzennych. Szczegółowe wartości wskaźników przedstawiono poniżej.

	WSKAŹNIK	STATYSTYKA Z	P-VALUE
MORAN I	0,158411074	5,018158987	5,22E-07
GEARY C	0,820733383	-4,84098953	6,46E-07
G GETIS I ORD	0,017834711	3,103738783	0,00095546

Wyniki analizy wskazują, że wskaźnik Morana I (**0,1584**) sugeruje umiarkowaną dodatnią autokorelację przestrzenną, co oznacza, że sąsiadujące powiaty wykazują pewne podobieństwo w ilości zebranych odpadów komunalnych. Statystyka Z (**5,018**) oraz bardzo niskie p-value (**5,22e-07**) pozwalają odrzucić hipotezę zerową, potwierdzając istotność globalnej zależności przestrzennej. Jednocześnie wskaźnik Geary'ego C (**0,8207**) wskazuje na negatywną autokorelację przestrzenną, czyli większe zróżnicowanie między sąsiadującymi powiatami, co również jest statystycznie istotne ($Z = -4,841$, $p\text{-value} = 6,46e-07$). Dodatkowo statystyka G Getisa i Orda (**0,0178**) identyfikuje obecność istotnych lokalnych skupisk (hot-

spotów), co potwierdzają $Z = 3,103$ oraz $p\text{-value} = 0,000955$. Wyniki te wskazują na istnienie zarówno globalnych wzorców przestrzennych, jak i lokalnych anomalii w rozkładzie odpadów komunalnych w powiatach.



Kartogram przedstawia lokalny wskaźnik Morana dla powiatów w Polsce, pokazując zależności przestrzenne w rozkładzie ilości odpadów komunalnych. Ciemnoczerwone obszary wskazują na lokalne skupiska powiatów, gdzie zarówno wartość zmiennej w danym powiecie, jak i w jego sąsiedztwie, są wysokie (dodatnia autokorelacja). Ciemnoniebieskie obszary pokazują miejsca o istotnych różnicach między wartością zmiennej w powiecie a jego sąsiedztwem (ujemna autokorelacja). Najwyraźniejsze skupiska dodatniej autokorelacji można zauważyć w okolicach większych miast - Warszawy, Krakowa oraz Wrocławia, co wskazuje na wyraźną koncentrację wysokich wartości w silnie zurbanizowanych regionach o wysokiej liczbie ludności. Z kolei obszary ujemnej autokorelacji występują w mniejszych miejscowościach lub regionach o bardziej zróżnicowanym sąsiedztwie, gdzie wartości zmiennej wyraźnie odbiegają od sąsiednich powiatów. Obszary o neutralnym kolorze sugerują brak istotnych zależności przestrzennych. Kartogram pozwala lepiej zrozumieć lokalne wzorce przestrzenne i wskazuje, które obszary wymagają szczególnej uwagi w analizie gospodarki odpadami, podkreślając zarówno koncentrację, jak i kontrasty przestrzenne. Pozwala to między innymi zidentyfikować regiony wymagające bardziej indywidualnego podejścia w planowaniu gospodarki odpadami, lub które wymagają optymalizacji procesów w tym zakresie.

5. Testy permutacyjne

W tej części dokonano analizy istotności autokorelacji przestrzennej, wykorzystując testy permutacyjne dla wskaźników Morana I, Geary'ego C oraz statystyki G Getisa i Orda.

	WSKAŹNIK	STATYSTYKA Z (Z PERMUTACJĄ)	P-VALUE (Z PERMUTACJĄ)
MORAN I	0,158411074	5,641948796	0,004
GEARY C	0,820733383	-	0,082
G GETIS I ORDA	0,017834711	-	0,016

W tej części przeprowadzono analizę istotności autokorelacji przestrzennej na poziomie powiatów, opierając się na testach permutacyjnych dla wskaźników Morana I, Geary'ego C oraz statystyki G Getisa i Orda. Wynik wskaźnika Morana I ($p = 0,004$) wykazał istotną dodatnią autokorelację przestrzenną, co sugeruje, że powiaty sąsiadujące mają zbliżone wartości pod względem ilości zebranych odpadów komunalnych. W przypadku wskaźnika Geary'ego C ($p = 0,082$) nie zaobserwowano istotnych różnic między sąsiadującymi powiatami. Z kolei statystyka G Getisa i Orda ($p = 0,016$) wskazuje na istnienie lokalnych skupisk (hot-spotów), w których obserwuje się wyraźnie wyższe lub niższe wartości zmiennej.

Podsumowując, testy permutacyjne potwierdziły obecność istotnej dodatniej autokorelacji przestrzennej w skali globalnej (Moran I) oraz lokalnych wzorców przestrzennych w postaci skupisk (G Getisa i Orda). Brak istotności wskaźnika Geary'ego C sugeruje natomiast, że zróżnicowanie wartości między sąsiadującymi powiatami nie jest wystarczająco wyraźne, aby uznać je za statystycznie istotne.

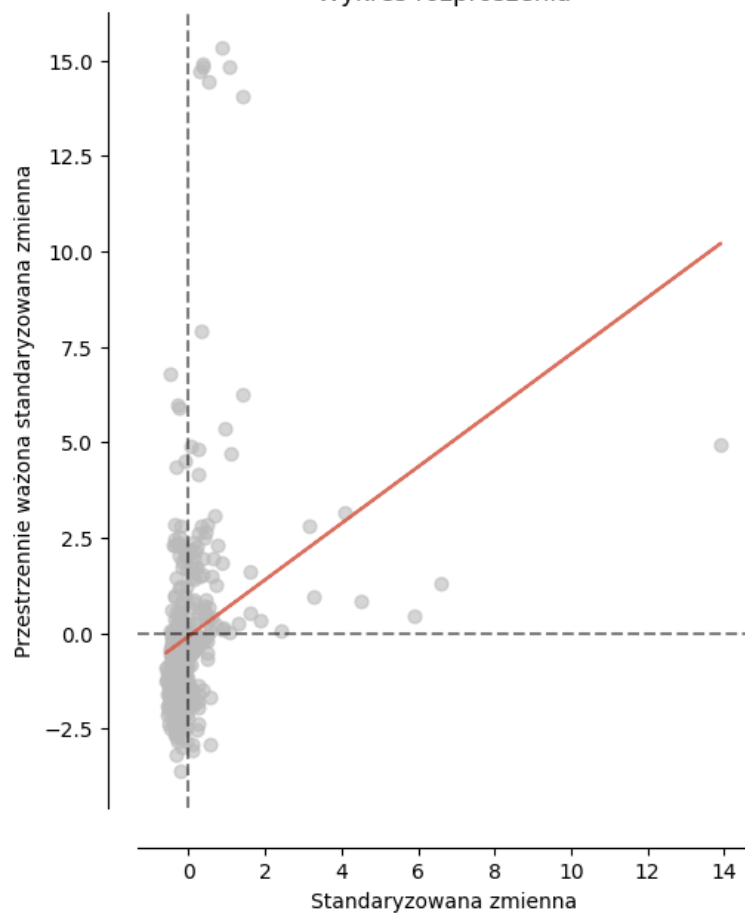
6. Wykres rozproszenia i korelogram

Wykres rozproszenia przedstawia wszystkie powiaty, wyznaczono dla nich współrzędne na osi OX i OY. Oś X reprezentuje różnicę między wartością badanej cechy w danym powiecie a jej średnią, co pozwala określić, czy wartość jest niska, przeciętna, czy wysoka w stosunku do pozostałych powiatów. Powiaty znajdujące się po lewej stronie wykresu charakteryzują się niższą wartością badanej cechy, natomiast te po prawej stronie mają wyższe wartości zmiennej. Powiaty zlokalizowane w pobliżu osi Y mają wartości cechy zbliżone do średniej.

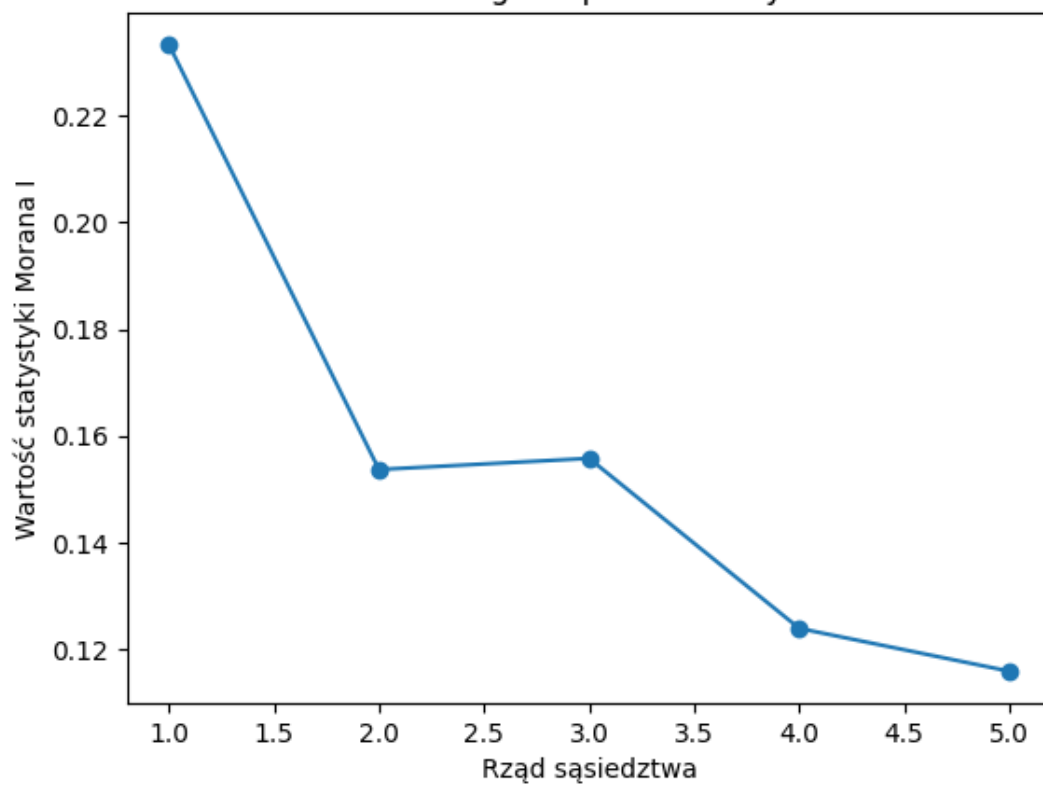
Oś Y przedstawia przestrzennie ważone wartości badanej zmiennej, uwzględniając wpływ sąsiednich powiatów. Wysokie wartości na osi Y wskazują, że sąsiednie powiaty również charakteryzują się wysokimi wartościami zmiennej, natomiast niskie wartości na osi Y oznaczają, że sąsiedztwo wykazuje niskie wartości cechy. Powiaty z wysokimi wartościami na obu osiach (X i Y) reprezentują obszary, gdzie badana cecha osiąga wysokie wartości zarówno lokalnie, jak i w sąsiedztwie (tzw. high-high). Z kolei wysokie wartości na osi X i niskie na osi Y wskazują na powiaty o wysokiej wartości cechy, otoczone regionami o wartościach niskich (high-low). Taka interpretacja umożliwia identyfikację zarówno lokalnych skupisk, jak i kontrastów przestrzennych między powiatami.

Jak możemy zauważyć na poniższym wykresie, większość punktów jest skoncentrowana w pobliżu osi Y, co sugeruje, że wiele powiatów ma wartości zmiennej bliskie średniej, a ich sąsiedztwo również charakteryzuje się podobnymi wartościami. Najwięcej punktów znajduje się w górnej prawej i dolnej lewej ćwiartce wykresu, co potwierdza istnienie lokalnych skupisk o podobnych wartościach (high-high i low-low). Z kolei niewielka liczba punktów w ćwiartkach high-low i low-high wskazuje na rzadkość występowania kontrastowych sąsiedztw, gdzie powiat o wysokiej wartości zmiennej sąsiedowałby z powiatami o niskich wartościach, lub odwrotnie. Wyniki podkreślają lokalny charakter zależności przestrzennych i wskazują, że wpływ sąsiedztwa na wartości zmiennej jest widoczny choć nie jest w pełni jednoznaczny.

Wykres rozproszenia



Korelogram przestrzenny



Korelogram przestrzenny dla wskaźnika Morana I przedstawia, jak zmienia się autokorelacja przestrzenna w zależności od liczby najbliższych sąsiadów (k). Najwyższą wartość wskaźnika autokorelacji przestrzennej, wynoszącą około 0,22, obserwujemy dla pierwszego rzędu sąsiedztwa, co wskazuje na wyraźne podobieństwo między bezpośrednio sąsiadującymi powiatami pod względem badanej zmiennej. Wraz ze wzrostem liczby sąsiadów, wartość wskaźnika systematycznie maleje, osiągając poziom około 0,12 dla pięciu najbliższych sąsiadów. Takie wyniki sugerują, że zależności przestrzenne mają przede wszystkim lokalny charakter – najbardziej widoczne są dla najbliższego otoczenia jednostek. W miarę uwzględniania kolejnych sąsiedztw wpływ sąsiadów staje się coraz słabszy, co wskazuje na zmniejszającą się siłę podobieństw między regionami przy większych dystansach. Korelogram podkreśla, że analizowana zmienna jest silniej powiązana w bezpośrednim sąsiedztwie. Najpewniej jest to związane z dużą liczbą odpadów zlokalizowanych w okolicach większych miast jak na przykład Warszawa, czy też dużej aglomeracji miejskiej w okolicach Katowic na Śląsku.

7. Analiza wyników dla różnych macierzy odległości

	Macierz wag	Moran I	Moran p-value	Geary C	Geary p-value	Getis-Ord G	Getis-Ord p-value
0	Macierz wag przestrzennych Queen	0.158411	0.003	0.820733	0.085	0.017835	0.011
1	Rook	0.158411	0.003	0.820733	0.071	0.017835	0.006
2	Macierz odległości z progiem	0.055788	0.029	1.010966	0.446	0.035076	0.006
3	Macierz KNN ($k=4$)	0.123869	0.004	0.847498	0.051	0.012619	0.012

Jak możemy zauważyć na podstawie wyników z powyższej tabeli, wybór macierzy wag ma istotny wpływ na wyniki analizy autokorelacji przestrzennej.

Tak jak wcześniej sugerowaliśmy, macierze Queen oraz Rook wykazują podobne wartości wskaźnika Morana I (0,1584, $p < 0,01$), co potwierdza umiarkowaną dodatnią autokorelację przestrzenną, sugerując, że oba typy sąsiedztwa mogą być zamienne w przypadku analizy powiatów. Statystyka Getis-Ord G (0,0178, $p < 0,05$) potwierdza istnienie lokalnych skupisk, co potwierdza przestrzenną koncentrację wysokich i niskich wartości w danych.

Dla macierzy odległościowej z progiem 50 km wartość wskaźnika Morana I wynosi 0,0558, a p-value na poziomie 0,029 umożliwia odrzucenie hipotezy zerowej o braku autokorelacji.

Wskazuje to, że występują zależności przestrzenne, choć są one słabsze i ograniczają się głównie do bliższego otoczenia. Wskaźnik Geary'ego C dla tej macierzy wynosi 1,0110, a wysoka p-value (0,446) wskazuje na brak istotnych różnic przestrzennych między sąsiadującymi jednostkami. Natomiast statystyka Getis-Ord G (0,0351, $p = 0,006$) sugeruje obecność lokalnych skupisk, co może oznaczać, że w pewnych obszarach występują skoncentrowane wzorce przestrzenne.

W przypadku macierzy KNN ($k=4$) wartość Morana I wynosi 0,1239, a niski poziom p-wartości (0,004) potwierdza istotność statystyczną. Wyniki wskazują na występowanie lokalnych zależności przestrzennych, w których kluczową rolę odgrywają najbliżsi sąsiedzi. Wartość Geary'ego C dla tej macierzy wynosi 0,8475, a p-value na poziomie 0,051 sugeruje umiarkowane różnice między sąsiadującymi jednostkami, choć na granicy istotności statystycznej. Statystyka Getis-Ord G (0,0126, $p = 0,012$) również wskazuje na obecność lokalnych skupisk wartości zmiennej. Zastosowanie tego rodzaju macierzy jest szczególnie korzystne w analizach, gdzie jednostki są rozmieszczone nierównomiernie, ponieważ zapewnia równą liczbę sąsiadów dla każdej jednostki.

Podsumowując, p-wartości pozwalają stwierdzić, że dla macierzy Queen, Rook i KNN zależności przestrzenne są istotne, natomiast macierz odległości wskazuje na słabszą, choć również statystycznie istotną autokorelację. Wyniki te pokazują, że wybór macierzy wag wpływa na postrzeganie siły zależności przestrzennych, a różne rodzaje macierzy pozwalają analizować zarówno globalne (macierz Queen, macierz Rook), jak i lokalne (macierz z progiem, macierz KNN) oddziaływania.

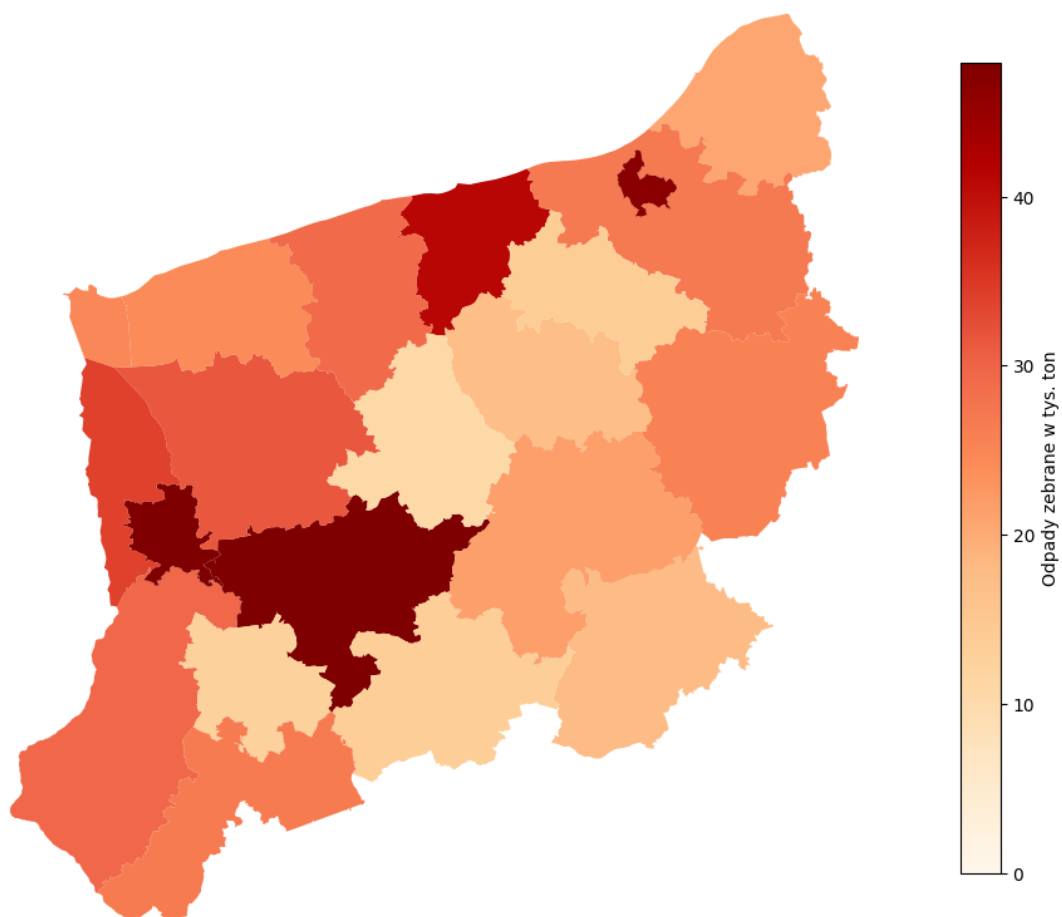
8. Analiza dwóch wybranych województw

8.1. Województwo Zachodniopomorskie

Rozkład danych jaki możemy zaobserwować na poniższej mapie charakteryzuje się wyraźnym skupieniem produkcji odpadów w południowo-zachodniej części województwa, gdzie znajduje się Szczecin – stolica województwa oraz główny ośrodek gospodarczy. Wybrzeże Bałtyku, choć turystyczne, wydaje się generować mniej odpadów, co może wynikać z mniejszej liczby stałych

mieszkańców, lecz sezonowo odpady mogą wzrastać. Dodatkowo możemy zauważyć, że zwiększona liczba odpadów jest widoczna w powiecie kołobrzeskim oraz koszalińskim, co ponownie łączy się z obecnością większych miast.

Rozkład surowych danych - zachodniopomorskie



8.1.1. Analiza autokorelacji przestrzennej

Analiza autokorelacji przestrzennej dla województwa zachodniopomorskiego, oparta na danych dotyczących ilości zebranych odpadów komunalnych, dostarcza następujących wyników:

1. Wskaźnik Morana I:

Wartość wskaźnika	0,0599
Statystyka z	0,7623
P-value	0,4459

Analiza przeprowadzona przy różnych macierzach wag przestrzennych daje następujące wyniki dla wskaźnika Morana I:

- Macierz Queen: Wartość wskaźnika Morana I wynosi 0,0599, a p-value wynosi 0,0670, co wskazuje na pewną tendencję do przestrzennej autokorelacji, jednak nie na poziomie wymaganym do odrzucenia hipotezy zerowej o braku autokorelacji przestrzennej.
- Macierz Rook: Wartość wskaźnika Morana I wynosi również 0,0599, z p-value równym 0,0790, co również sugeruje potencjalną przestrzenną autokorelację, ale nie na poziomie istotnym statystycznie.
- Macierz odległości z progiem: Wskaźnik Morana I jest bardzo niski i wynosi 0,0088, przy braku istotności statystycznej. Wartości te nie wskazują na istnienie istotnej autokorelacji przestrzennej.
- Macierz KNN (k = 4): Wartość wskaźnika Morana I wynosi -0,0245, a p-value również nie jest istotne statystycznie, co sugeruje brak przestrzennej autokorelacji w analizowanych danych.

2. Wskaźnik Geary'ego C:

Wartość wskaźnika	0,7742
Statystyka z	-1,3979
P-value	0,0811

W przypadku **wskaźnika Geary'ego C**, uzyskano wartość 0,7742 przy statystyce z równej - 1,3979 oraz p-value wynoszącym 0,0811. Chociaż wynik wskazuje na możliwość istnienia dodatniej autokorelacji przestrzennej, wysoki poziom p-value nie pozwala na stwierdzenie istotnej zależności. Innymi słowy, rozkład odpadów komunalnych w poszczególnych jednostkach administracyjnych nie jest dostatecznie podobny, aby odrzucić hipotezę zerową o braku autokorelacji.

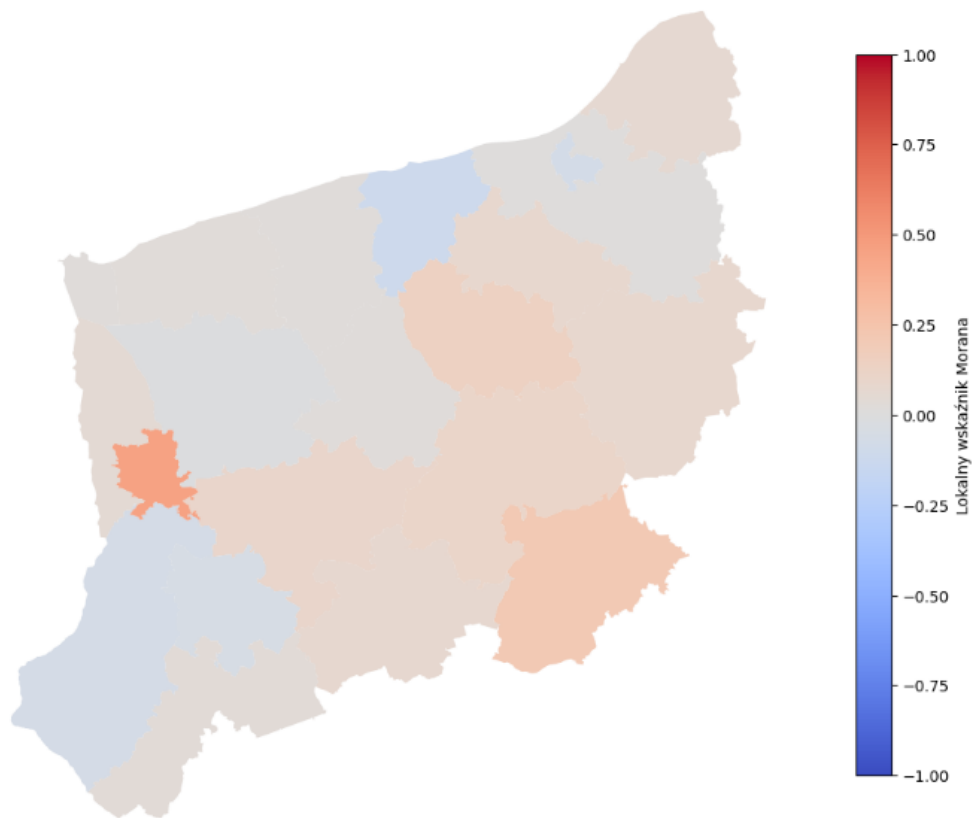
3. Statystyka G Getisa i Orda:

Wartość wskaźnika	0,2373
Statystyka z	0,9192
P-value	0,0811

Dodatkowo, analiza **statystyki G Getisa i Orda** dostarczyła wartość wskaźnika równą 0,2373 oraz p-value wynoszące 0,2063. Statystyka ta, wskazująca na możliwość występowania lokalnych skupisk wysokich lub niskich wartości, nie wskazała statystycznie istotnych lokalnych skupisk, co oznacza, że nie zaobserwowano żadnych przestrzennych zgrupowań wysokich lub niskich wartości odpadów w analizowanym regionie.

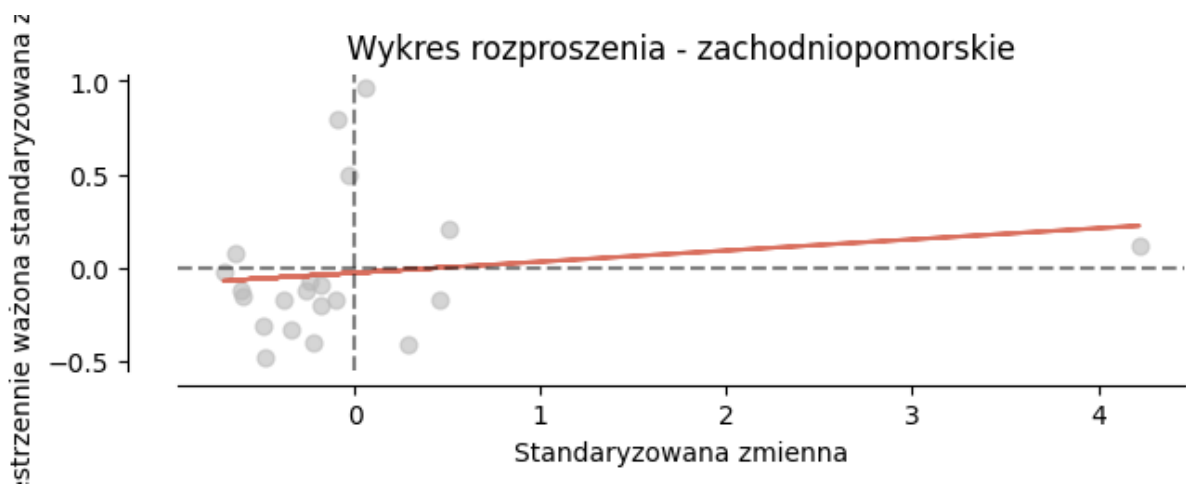
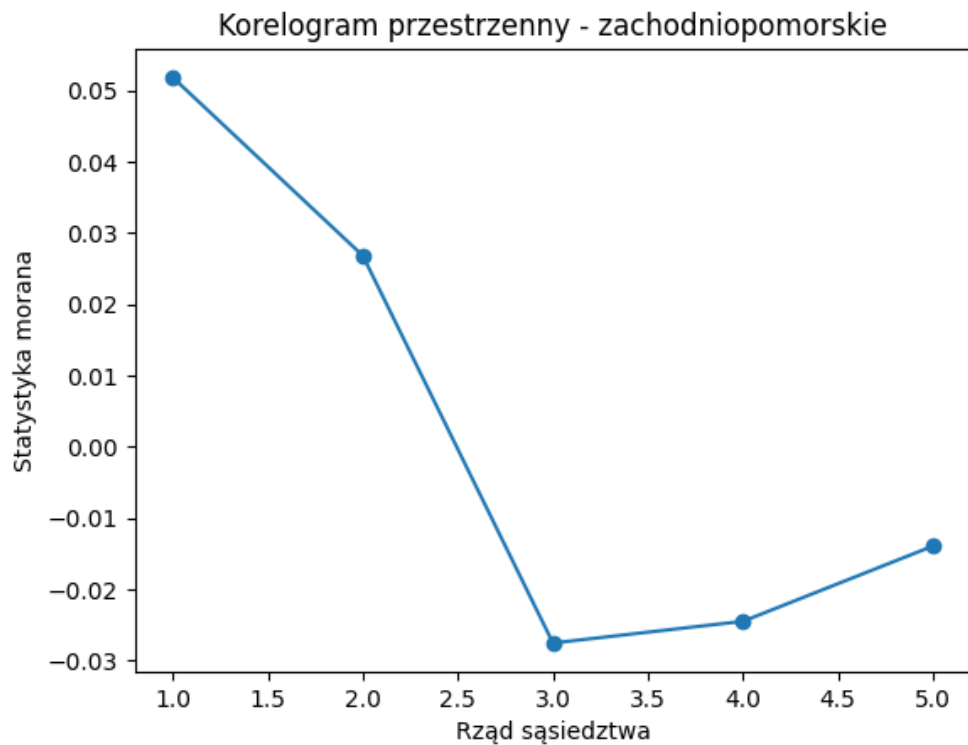
Przeprowadzone analizy nie wykazały istotnej autokorelacji przestrzennej w rozkładzie ilości zebranych odpadów komunalnych w województwie zachodniopomorskim. Zarówno globalne miary (wskaźnik Morana I), jak i lokalne (wskaźnik Geary'ego C oraz statystyka G Getisa i Orda) nie wskazują na statystycznie istotne zależności przestrzenne. Oznacza to, że ilość odpadów w poszczególnych powiatach nie jest istotnie powiązana z ilością odpadów w sąsiednich powiatach. Na poniższej mapie przedstawiającej lokalny wskaźnik Morana dla powiatów, możemy zaobserwować wystąpienie większej wartości w zachodniej części województwa, co związane jest z obecnością Szczecina. Potwierdza to występowanie lokalnych skupisk, mimo że globalny wskaźnik Morana I nie wskazuje na istotną autokorelację przestrzenną. Niemniej jednak widzimy, że wartości w całym województwie są bardzo niskie co faktycznie potwierdza wcześniej wysnute wnioski.

Lokalny wskaźnik Morana - zachodniopomorskie



8.1.2. Wykres rozproszenia i korelogram

Na poniższym korelogramie przestrzennym można zauważyć, że wartość wskaźnika Morana I maleje wraz ze wzrostem liczby najbliższych sąsiadów (k). Dla pierwszych dwóch poziomów sąsiedztwa wskaźnik przyjmuje dodatnie wartości, wskazując na słabą dodatnią autokorelację przestrzenną. Po osiągnięciu wartości $k = 3$ wskaźnik Morana I przechodzi na ujemną stronę, co oznacza brak autokorelacji lub nawet potencjalne różnice przestrzenne w dalszych sąsiedztwach. To sugeruje, że przestrzenne podobieństwo jest ograniczone do lokalnych relacji.



Wykres rozproszenia przedstawia zależność między standardyzowaną wartością zmiennej (ilością odpadów komunalnych) a jej przestrzennie ważoną wartością dla województwa zachodniopomorskiego. Nachylenie czerwonej linii regresji jest dodatnie, lecz niewielkie, co wskazuje na bardzo słabą dodatnią autokorelację przestrzenną. Rozkład punktów na wykresie, skoncentrowany głównie wokół osi poziomej i pionowej, sugeruje brak wyraźnych zależności przestrzennych na poziomie globalnym.

Powyższe wnioski są spójne z wartością p-value dla wskaźnika Morana I (0,4459), co nie pozwala na odrzucenie hipotezy zerowej o braku istotnej autokorelacji przestrzennej. Wyniki

wskazują, że w analizowanym regionie przestrzenne wzorce są słabe, a wpływ sąsiedztwa na badane wartości jest nieznaczny.

8.1.3 Analiza wyników dla różnych macierzy odległości

	Macierz wag	Moran I	Moran p-value	Geary C	Geary p-value	Getis-Ord G	Getis-Ord p-value
0	Macierz wag przestrzennych Queen	0.059932	0.077	0.774231	0.144	0.237263	0.204
1	Macierz wag przestrzennych Rook	0.059932	0.071	0.774231	0.149	0.237263	0.177
2	Macierz odległości z progiem	0.008818	0.154	0.999732	0.482	0.247446	0.217
3	Macierz KNN (k=4)	-0.024508	0.350	1.085283	0.307	0.218156	0.152

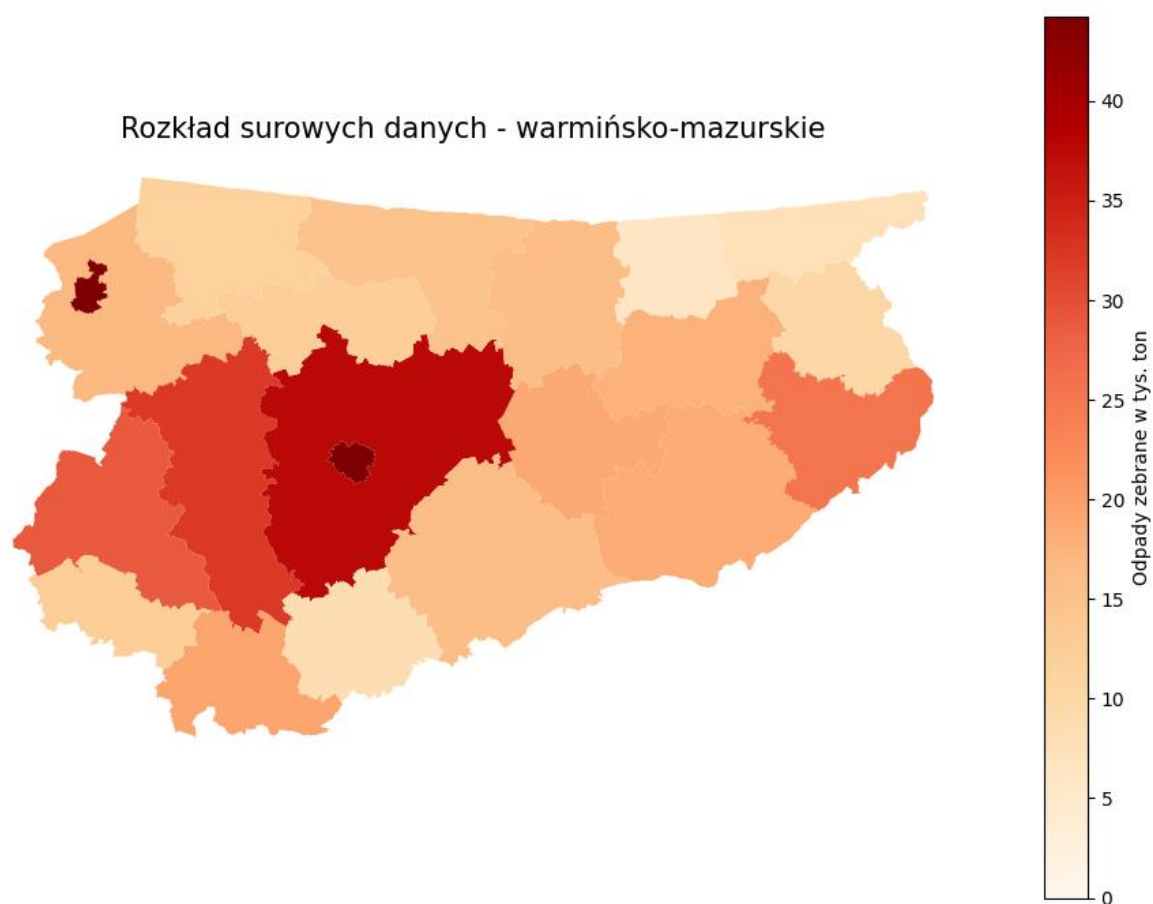
Podobnie jak w przypadku analizy dla całej Polski, można zauważyć podobieństwo wyników macierzy wag Queen i Rook. Jak wynika z analizy przedstawionych wyników, p-value dla większości macierzy wag wskazują na brak istotności statystycznej wyników w zakresie autokorelacji przestrzennej. Wyjątek stanowią jedynie macierz Queen i Rook, gdzie p-value dla statystyki Morana są bliskie granicy istotności (odpowiednio 0,077 i 0,071). Wartości wskaźnika Morana I dla tych macierzy są niskie (0,0599), co sugeruje bardzo słabą dodatnią autokorelację przestrzenną, jednak nie wystarczająco istotną, by odrzucić hipotezę zerową przy typowym poziomie istotności 0,05.

W przypadku macierzy odległości z progiem wartość Morana I wynosi 0,0088, a p-value 0,154 wskazuje na brak podstaw do odrzucenia hipotezy zerowej. Wartości statystyki Geary'ego C (0,9997, p-value = 0,482) oraz statystyki Getis-Ord G (0,2474, p-value = 0,217) również sugerują brak wyraźnych wzorców przestrzennych i wskazują na losowe rozmieszczenie zmiennej w przestrzeni. Wyniki te mogą świadczyć o tym, że wpływ przestrzenny jest ograniczony lub niemal nieistniejący.

Macierz KNN (k=4) wyróżnia się ujemną wartością wskaźnika Morana I (-0,0245), co wskazuje na potencjalnie ujemną autokorelację, jednak p-value (0,350) jest daleka od istotności statystycznej. Statystyka Geary'ego C (1,0853, p-value = 0,307) oraz statystyka Getis-Ord G (0,2182, p-value = 0,152) również wskazują na brak istotnych wzorców przestrzennych w lokalnych zależnościach.

Podsumowując, analiza wyników wskazuje na brak istotnej autokorelacji przestrzennej dla większości zastosowanych macierzy wag. Macierze Queen i Rook wykazują jedynie marginalną istotność, co może sugerować występowanie słabych zależności przestrzennych, choć nie można tego stwierdzić z wystarczającą pewnością statystyczną.

8.2. Województwo Warmińsko-Mazurskie



Centralne obszary (Olsztyn i okolice) mają największą koncentrację odpadów, co wynika z dużej liczby mieszkańców oraz rozwoju miejskiego. Południowo-wschodnie części województwa cechują się niższymi wartościami – obszary te są bardziej wiejskie i mniej zaludnione. W północno-wschodnich rejonach (bliżej granicy z Rosją) obserwujemy obszary o niższych wartościach, co może wynikać z większej izolacji tych terenów.

8.2.1. Analiza autokorelacji przestrzennej

1. Wskaźnik Morana I:

Wartość wskaźnika	0,2253
Statystyka z	1,8344
P-value	0,0666

Przeprowadzona analiza globalnej autokorelacji przestrzennej za pomocą wskaźnika Morana I w województwie warmińsko-mazurskim wskazuje na wartość 0,2253 przy poziomie istotności (p-value) wynoszącym 0,0666. Otrzymane wyniki sugerują, że nie można odrzucić hipotezy zerowej o braku istotnej autokorelacji przestrzennej na typowym poziomie istotności $\alpha=0,05$. Wartość wskaźnika wskazuje jednak na potencjalnie słabą pozytywną autokorelację przestrzenną, co może oznaczać, że sąsiadujące jednostki administracyjne wykazują umiarkowaną tendencję do podobnych wartości w zakresie zbiórki odpadów komunalnych.

Analiza przeprowadzona przy różnych macierzach wag przestrzennych daje następujące wyniki dla wskaźnika Morana I:

- **Macierz Queen:** Wartość wskaźnika Morana I wynosi 0,2253, a p-value z permutacji to 0,0410, co wskazuje na istotną autokorelację przestrzenną na poziomie $\alpha=0,05$.
- **Macierz Rook:** Wartość wskaźnika Morana I również wynosi 0,2253, z p-value równym 0,0420, co potwierdza istotność autokorelacji przestrzennej.
- **Macierz odległości z progiem:** Wskaźnik Morana I jest znacznie niższy i wynosi 0,0161, przy p-value równym 0,3280, co nie wskazuje na istnienie autokorelacji przestrzennej.
- **Macierz KNN (k = 4):** Wskaźnik Morana I wynosi 0,1179, a p-value to 0,0920, co sugeruje brak istotnej autokorelacji przestrzennej na typowym poziomie istotności.

Najbardziej przekonujące wyniki uzyskano dla macierzy Queen i Rook, wskazujących na umiarkowaną, pozytywną autokorelację przestrzenną.

2. Wskaźnik Geary'ego C:

Wartość wskaźnika	0,5237
-------------------	--------

Statystyka z	-2,8901
P-value	0,0019

Wskaźnik Geary'ego C obliczony dla macierzy wag Queen wynosi 0,5237, co jest wartością istotnie niższą od 1. P-value wynosi 0,0019, co pozwala na odrzucenie hipotezy zerowej o braku autokorelacji przestrzennej na poziomie istotności $\alpha=0,05$. Otrzymany wynik wskazuje na występowanie istotnej dodatniej autokorelacji przestrzennej, co oznacza, że sąsiadujące jednostki administracyjne mają bardzo podobne wartości zmiennej dotyczącej zbiórki odpadów komunalnych.

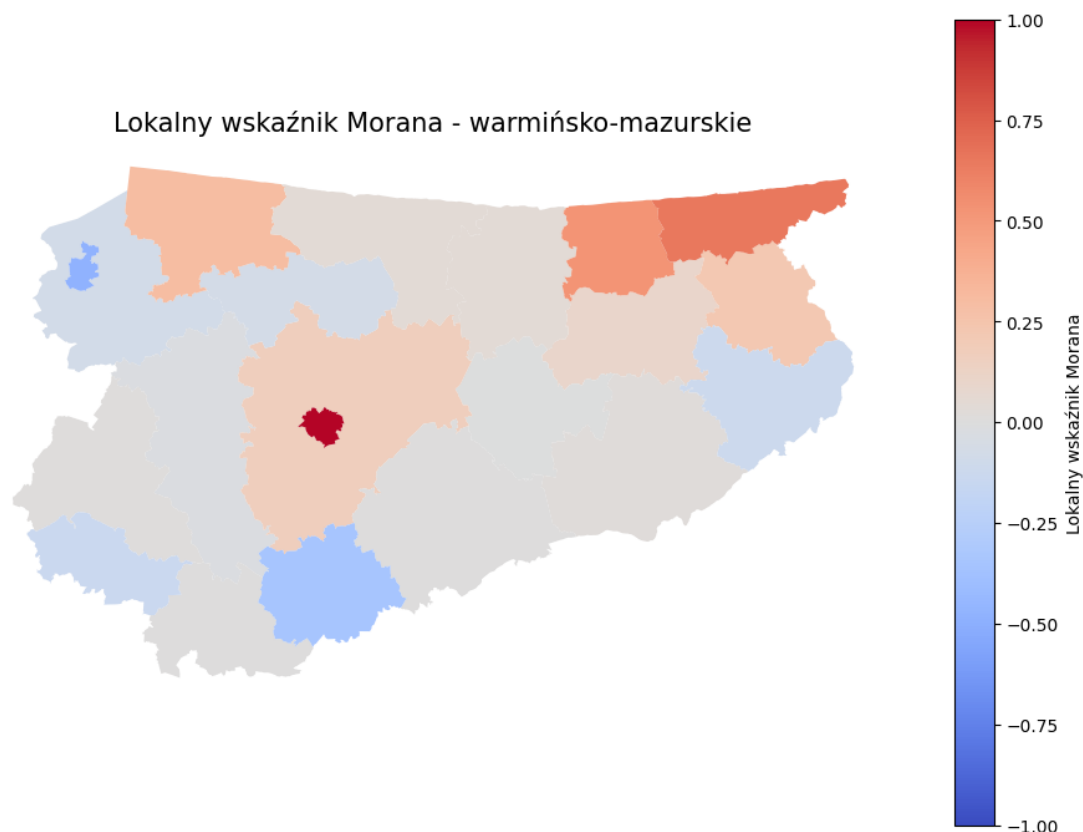
3. Statystyka G Getisa i Orda:

Wartość wskaźnika	0,1891
Statystyka z	-0,2377
P-value	0,406

Analiza lokalnej autokorelacji przestrzennej za pomocą statystyki G Getisa i Orda wskazuje na wartość wskaźnika równą 0,1891, z wartością statystyki z wynoszącą -0,2377. Otrzymane p-value wynosi 0,406, co oznacza brak podstaw do odrzucenia hipotezy zerowej o braku istotnej autokorelacji przestrzennej na typowym poziomie istotności $\alpha=0,05$.

Wyniki te wskazują, że w badanym układzie przestrzennym brak jest dowodów na istnienie istotnych lokalnych zależności przestrzennych w wartościach analizowanej zmiennej. Statystyka G Getisa i Orda nie wykazała obecności skupisk wysokich lub niskich wartości w województwie warmińsko-mazurskim.

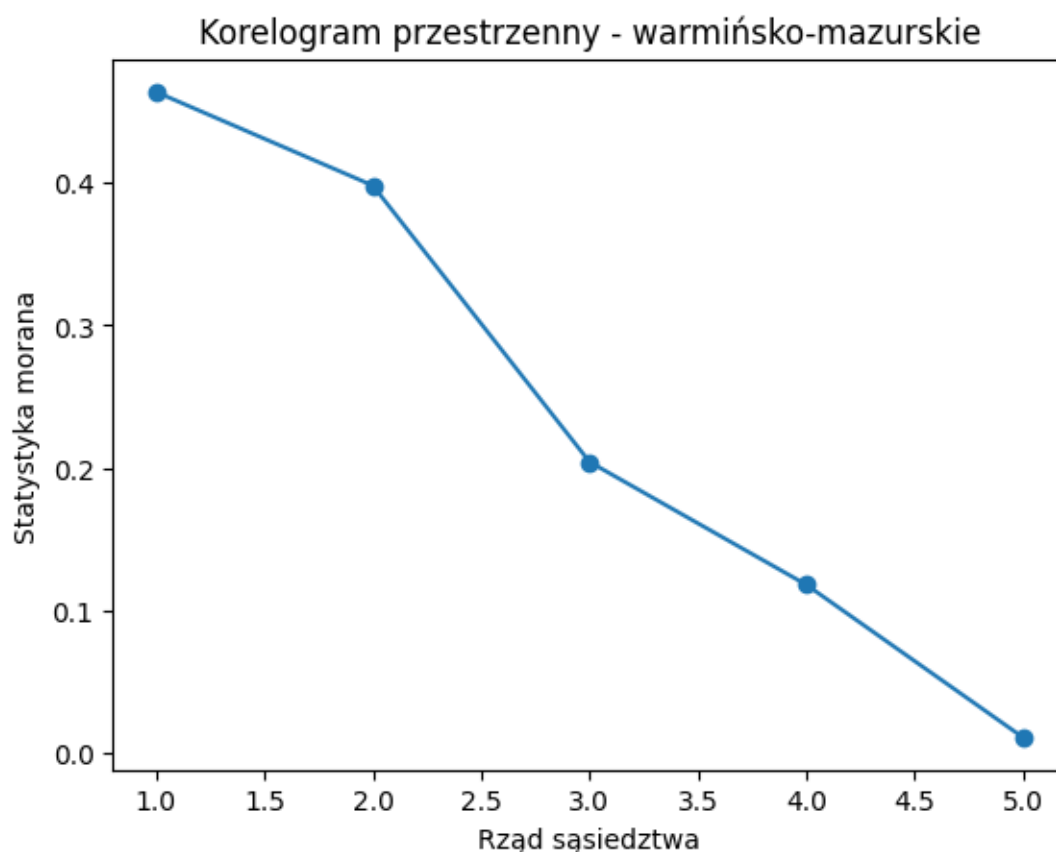
Analiza autokorelacji przestrzennej dla województwa warmińsko-mazurskiego wskazuje na brak istotnej globalnej autokorelacji według wskaźnika Morana I, ale wskaźnik Geary'ego C potwierdza statystycznie istotne różnice między sąsiadującymi powiatami. Brak istotnych wyników statystyki G Getisa i Orda sugeruje brak wyraźnych lokalnych skupisk, co podkreśla przestrzenną heterogeniczność województwa.



Mapa lokalnego wskaźnika Morana dla województwa warmińsko-mazurskiego wskazuje na wyraźne obszary dodatniej autokorelacji w północno-wschodnich powiatach (wysokie wartości sąsiadują z wysokimi) oraz ujemnej autokorelacji w południowo-zachodnich regionach (duże różnice między sąsiadami). Statystycznie istotny wskaźnik Geary'ego C potwierdza przestrzenne zróżnicowanie między powiatami, natomiast brak istotnych wyników statystyki G Getisa i Orda sugeruje brak wyraźnych, statystycznie istotnych lokalnych skupisk.

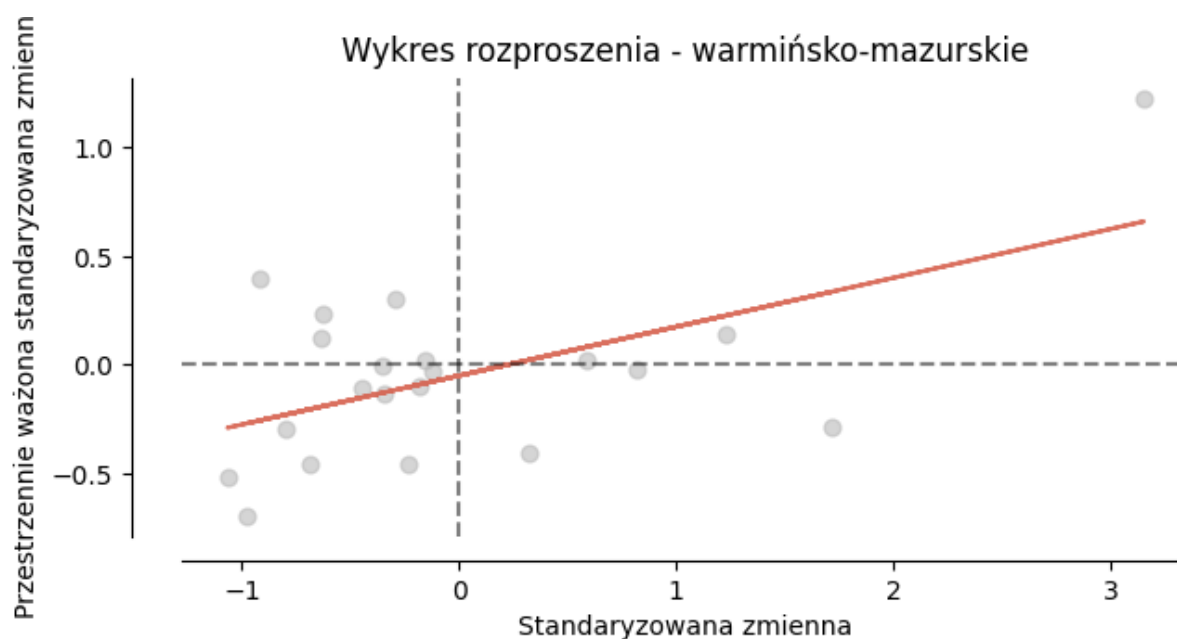
8.2.2. Wykres rozproszenia i korelogram

Korelogram pokazuje, że wartość wskaźnika Morana I systematycznie maleje wraz ze wzrostem liczby najbliższych sąsiadów (k), od umiarkowanej dodatniej autokorelacji dla $k = 1$ (0,4) do wartości bliskiej zeru dla $k = 5$. Wskazuje to, że przestrzenne podobieństwo w ilości odpadów jest zauważalne jedynie w najbliższym sąsiedztwie.



Poniższy wykres rozproszenia przedstawia zależność między standaryzowaną ilością odpadów komunalnych w powiatach dla województwa warmińsko-mazurskiego, a ich przestrzennie ważoną średnią w tym obszarze. Nachylenie czerwonej linii regresji jest dodatnie, co wskazuje na dodatnią autokorelację przestrzenną. Większość punktów skupia się w pobliżu środka układu współrzędnych, co sugeruje, że większość powiatów ma ilość odpadów zbliżoną do średniej regionalnej, a ich sąsiedztwa również charakteryzują się podobnymi wartościami.

Wskaźnik Morana I (0,2253) wskazuje na umiarkowaną zależność przestrzenną, jednak p-value (0,0666) nie pozwala na odrzucenie hipotezy zerowej o braku istotnej autokorelacji przestrzennej. Wnioski z analizy potwierdzają, że w badanym regionie zależności przestrzenne w zakresie ilości odpadów komunalnych są słabe i nie mają wyraźnego charakteru globalnego.



8.2.3 Analiza wyników dla różnych macierzy odległości

	Macierz wag	Moran I	Moran p-value	Geary C	Geary p-value	Getis-Ord G	Getis-Ord p-value
0	Macierz wag przestrzennych Queen	0.225338	0.028	0.523718	0.004	0.189092	0.422
1	Macierz wag przestrzennych Rook	0.225338	0.050	0.523718	0.005	0.189092	0.453
2	Macierz odległości z progiem	0.016127	0.304	0.956880	0.353	0.199964	0.184
3	Macierz KNN (k=4)	0.117914	0.092	0.879355	0.185	0.226826	0.042

Dla macierzy Queen wartość wskaźnika Morana I wynosi 0,2253, a p-value równe 0,028 pozwala na odrzucenie hipotezy zerowej o braku autokorelacji przestrzennej. Wskazuje to na fakt, że powiaty o podobnych wartościach zmiennej grupują się przestrzennie. Statystyka Geary'ego C wynosi 0,5237 z p-value równym 0,004, co również potwierdza istotną dodatnią autokorelację. Statystyka Getis-Ord G (0,1891, p-value = 0,422) sugeruje brak istotnych lokalnych skupisk wartości zmiennej, co oznacza, że przestrzenne wzorce w tym przypadku nie są wyraźnie widoczne. Dla macierzy Rook możemy wyciągnąć podobne wnioski, różnicą tutaj jest statystyka Morana I gdzie p-value znajduje się na granicy istotności, mimo tego wnioski pozostają niezmiennie.

Macierz odległości z progiem daje znacznie niższe wartości wskaźnika Morana I (0,0161, p-value = 0,304), co wskazuje na brak istotnych zależności przestrzennych w większym zasięgu. Statystyka Geary'ego C (0,9569, p-value = 0,353) oraz statystyka Getis-Ord G (0,1999, p-value = 0,184) również nie wykazują żadnych znaczących wzorców przestrzennych, co sugeruje

bardziej losowy charakter rozmieszczenia zmiennej. Różnice to mogą wynikać z odległości pomiędzy centroidami województw.

Dla macierzy KNN ($k=4$) wartość wskaźnika Morana I wynosi 0,1179, a p-value równe 0,092 , co wskazuje na możliwość słabej dodatniej autokorelacji. Statystyka Geary'ego C (0,8794, p-value = 0,185) oraz statystyka Getis-Ord G (0,2268, p-value = 0,042) sugerują potencjalne istnienie lokalnych skupisk, przy czym statystyka Getis-Ord G dla tej macierzy osiąga poziom istotności statystycznej. Różnice w tych wynikach w stosunku do macierzy odległości ponownie mogą sugerować duże odległości pomiędzy centroidami.

9. Porównanie uzyskanych wyników

Wskaźnik Morana I

- **Polska:** Wskaźnik wynosi **0,1584**, co sugeruje umiarkowaną dodatnią autokorelację przestrzenną. Wynik jest istotny statystycznie (p-value < 0,05), co oznacza, że sąsiadujące powiaty mają podobne wartości.
- **Zachodniopomorskie:** Wartość wskaźnika jest bardzo niska (**0,0599**) i nieistotna statystycznie (p-value = 0,4459). Brak dowodów na autokorelację przestrzenną w tym województwie.
- **Warmińsko-Mazurskie:** Wskaźnik wynosi **0,2253**, co sugeruje umiarkowaną dodatnią autokorelację, ale wynik nie jest istotny statystycznie (p-value = 0,0666).

MORAN I	POLSKA	WOJEWÓDZTWO ZACHODNIOPOMORSKIE	WOJEWÓDZTWO WARMINSKO-MAZURSKIE
WSKAŹNIK	0,158411074	0,0599	0,2253
STATYSTYKA Z	5,018158987	0,7623	1,8344
P-VALUE	5,22E-07	0,4459	0,0666

Wskaźnik Geary'ego C

- **Polska:** Wartość **0,8207** wskazuje na ujemną autokorelację przestrzenną (zróznicowanie między sąsiadującymi powiatami). Wynik jest istotny statystycznie (p-value < 0,05).

- **Zachodniopomorskie:** Wartość **0,7742** sugeruje umiarkowaną ujemną autokorelację, jednak wynik jest nieistotny statystycznie (p-value = 0,0811).
- **Warmińsko-Mazurskie:** Wskaźnik wynosi **0,5237**, co oznacza silniejszą ujemną autokorelację przestrzenną. Wynik jest istotny statystycznie (p-value = 0,0019).

GEARY C	POLSKA	WOJEWÓDZTWO ZACHODNIOPOMORSKIE	WOJEWÓDZTWO WARMINSKO-MAZURSKIE
WSKAŹNIK	0,820733383	0,7742	0,5237
STATYSTYKA Z	5,018158987	-1,3979	-2,8901
P-VALUE	6,46E-07	0,0811	0,0019

Statystyka G Getisa i Orda

- **Polska:** Wartość **0,0178** wskazuje na obecność lokalnych skupisk wysokich/niskich wartości. Wynik jest istotny statystycznie (p-value < 0,05).
- **Zachodniopomorskie:** Wartość **0,2373** jest wyższa, ale wynik nie jest istotny statystycznie (p-value = 0,2063). Brak wyraźnych skupisk w województwie.
- **Warmińsko-Mazurskie:** Wartość **0,1891** nie wskazuje na istotne lokalne skupiska (p-value = 0,406).

G GETIS ORDA	POLSKA	WOJEWÓDZTWO ZACHODNIOPOMORSKIE	WOJEWÓDZTWO WARMINSKO-MAZURSKIE
WSKAŹNIK	0,017834711	0,2373	0,1891
STATYSTYKA Z	3,103738783	0,8192	-0,2377
P-VALUE	9,55E-04	0,2063	0,406

W Polsce obserwuje się umiarkowaną globalną autokorelację przestrzenną (Morana I), ujemną lokalną autokorelację (Geary'ego C) oraz wyraźne lokalne skupiska (G Getisa i Orda). W skali kraju przestrzenne zależności są silniejsze i bardziej wyraźne, co sugeruje, że polityka gospodarki odpadami powinna uwzględniać koncentrację w regionach miejskich. W województwie zachodniopomorskim nie stwierdzono istotnej autokorelacji zarówno globalnej, jak i lokalnej, co sugeruje, że region charakteryzuje się równomiernym rozkładem danych. W województwie warmińsko-mazurskim występują silniejsze różnice przestrzenne (ujemna autokorelacja Geary'ego C), jednak brak tam istotnych lokalnych skupisk, co wskazuje na większą heterogeniczność przestrzenną tego obszaru. Wyniki sugerują, że rozkład ilości

odpadów jest bardziej przestrzennie powiązany na poziomie całego kraju niż na poziomie województw. Warmińsko-Mazurskie wyróżnia się większymi kontrastami między powiatami niż Zachodniopomorskie.

10. Porównanie wskaźników autokorelacji dla powiatów i województw (wcześniejszy projekt)

Wyniki analizy statystycznej dla wskaźnika Morana I wskazują na umiarkowaną dodatnią autokorelację przestrzenną dla powiatów, co potwierdza wartość 0,1584 ($p\text{-value} < 0,05$). Natomiast dla województw wskaźnik ten wynosi -0,2949 i nie jest istotny statystycznie ($p\text{-value} = 0,1249$), co sugeruje brak wyraźnej autokorelacji przestrzennej na tym poziomie. Różnica ta wskazuje na większą zależność przestrzenną między sąsiadującymi jednostkami na poziomie powiatów niż województw.

Wskaźnik Geary'ego C dla powiatów wynosi 0,8207, co wskazuje na umiarkowaną ujemną autokorelację przestrzenną i jest istotny statystycznie ($p\text{-value} < 0,05$). Dla województw wartość tego wskaźnika wynosi 1,2736, co sugeruje tendencję do dodatniej autokorelacji, przy $p\text{-value} = 0,034$. Wyniki te wskazują na większe różnice między sąsiednimi jednostkami przestrzennymi na poziomie powiatów w porównaniu z województwami.

Wskaźnik Getisa i Orda dla powiatów wynosi 0,0178 i jest istotny statystycznie ($p\text{-value} < 0,05$), co wskazuje na obecność lokalnych skupisk wysokich/niskich wartości. W przypadku województw wskaźnik ten wynosi 0,2903, lecz nie jest istotny statystycznie ($p\text{-value} = 0,406$), co oznacza brak wyraźnych lokalnych wzorców.

Podsumowując, na poziomie powiatów występuje większa autokorelacja przestrzenna oraz obecność lokalnych skupisk w porównaniu z województwami. Wyniki te sugerują, że zależności przestrzenne między sąsiadującymi jednostkami są bardziej wyraźne na poziomie lokalnym niż regionalnym.

DLA WOJEWÓDZTW:

	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

DLA POWIATÓW:

	WSKAŹNIK	STATYSTYKA Z	P-VALUE
MORAN I	0,158411074	5,018158987	5,22E-07
GEARY C	0,820733383	-4,84098953	6,46E-07
G GETIS I ORD	0,017834711	3,103738783	0,00095546

11. Podsumowanie

Projekt miał na celu przeprowadzenie analizy autokorelacji przestrzennej ilości odpadów zebranych w powiatach w Polsce oraz w wybranych województwach: Zachodniopomorskim i Warmińsko-Mazurskim. W tym celu wykorzystano różne wskaźniki statystyczne, takie jak Morana I, Geary'ego C i Getisa-Orda G, oraz różne macierze wag przestrzennych, w tym Queen, Rook, KNN i macierz odległościową z progiem.

Analiza wykazała, że w skali całego kraju występuje umiarkowana dodatnia autokorelacja przestrzenna, co potwierdzają wartości wskaźnika Morana I ($p\text{-value} < 0,05$). Sąsiadujące powiaty wykazują podobne wartości zmiennej, co wskazuje na istotny wpływ przestrzennego rozmieszczenia na ilość odpadów. Wskaźnik Geary'ego C wskazuje na istotną ujemną autokorelację, co sugeruje znaczące różnice między sąsiadującymi powiatami, szczególnie w kontekście kontrastu między regionami miejskimi i wiejskimi. Statystyka Getisa-Orda G potwierdziła obecność lokalnych skupisk odpadów, które można przypisać głównie większym miastom.

Na poziomie województw wyniki były mniej wyraźne. W województwie Zachodniopomorskim nie zaobserwowano istotnych wzorców przestrzennych, co wskazuje na bardziej równomierny

rozkład odpadów w tym regionie. W województwie Warmińsko-Mazurskim zauważono większe różnice między sąsiadującymi powiatami, co potwierdził wskaźnik Geary'ego C ($p\text{-value} < 0,05$). Brak istotnych wyników dla statystyki Getisa-Orda G w obu województwach sugeruje, że lokalne skupiska są tutaj mniej wyraźne niż na poziomie kraju.

Podsumowując, przestrzenne wzorce w rozkładzie odpadów są bardziej widoczne w skali kraju, szczególnie między regionami miejskimi i wiejskimi. Na poziomie województw zależności przestrzenne są mniej istotne i mogą wynikać z lokalnych uwarunkowań gospodarczych. Badania podkreślają znaczenie doboru macierzy wag i wskaźników statystycznych oraz sugerują potrzebę uwzględnienia różnorodności przestrzennej w polityce gospodarki odpadami na poziomie kraju i regionów.