

**Analiza wielowymiarowa działalności
człowieka w Polsce oraz próba identyfikacji
województwa, którego krajobraz jest
najmniej naruszony społecznie i
infrastrukturalnie na podstawie danych
statystycznych z 2021 roku**



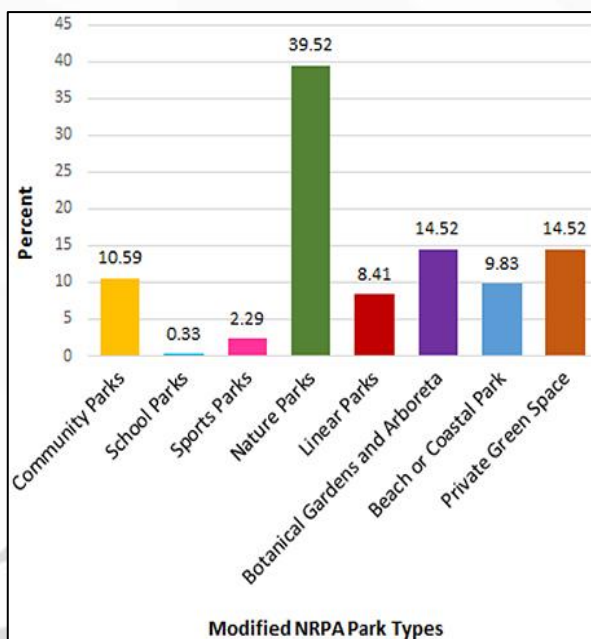
Wstęp

Zanieczyszczenie powietrza, brak zieleni i terenów rekreacyjnych, hałas, brak prywatności to tylko kilka czynników z którymi mogą mierzyć się mieszkańcy wysoce zurbanizowanych i zaludnionych terenów polski. W 2015 roku w czasopiśmie naukowym *International Journal of Environmental Research and Public Health* opublikowano badanie autorstwa R. U. Alok oraz J. V. Sunil pt. *Urbanization and mental health: a review of the literature and a public health response* [1], w którym przeanalizowano wpływ urbanizacji na zdrowie psychiczne i fizyczne, w tym na takie zagadnienia jak stres, niepokój, depresja i bezsenność, a także na zdrowie fizyczne, takie jak choroby serca, cukrzyca i otyłość. Autorzy stwierdzili, że życie w mieście i ośrodkach zurbanizowanych wiąże się z wyższym ryzykiem wystąpienia problemów zdrowotnych, takich jak te opisane powyżej, a także że urbanizacja może mieć negatywny wpływ na jakość życia i dobre samopoczucie. Wnioski te opierają się na analizie wielu badań przeprowadzonych na temat zdrowia i dobrego samopoczucia w mieście i ośrodkach zurbanizowanych. Również w badaniu opublikowanym w *Journal of Environmental Psychology* w 2013 roku, pod kierunkiem Dr. Gilesy Corey'ego z University of Utah o tytule *The Psychological Benefits of Urban Nature: Exploring the Value of Community Gardens* [2] wykazano, że mieszkańcy miast są narażeni na wyższy poziom stresu niż mieszkańcy wsi. W publikacji zwrócono również uwagę na to, że poziom hałasu, zanieczyszczenia powietrza i brak terenów zielonych przyczyniają się do większego stresu i pogorszenia jakości życia w miastach.

Pierwotna subiektywna obserwacja poparta między innymi powyższymi badaniami skłoniła do przyjrzenia się bliżej temu zagadnieniu, analizy tego zjawiska i próby znalezienia remedium na wymienione wcześniej bolączki współczesnego świata, jednak społeczeństwo już znalazło sposób jak sobie z tym radzić.

W badaniu *The Desire for Nature: Examining the Pull of the Natural World* autorstwa Dr. Christophera Neale'a i Dr. Davida Strouda z University of Derby opublikowanego w *Journal of Environmental Psychology* w 2019 roku [3], uczestniczyli dorośli mieszkańcy miast i ich potrzeba kontaktu z naturą. Wyniki pokazały, że ludzie coraz częściej odczuwają potrzebę kontaktu z naturą i odcinania się od codzienności w miejskim środowisku (co potwierdza również wcześniejsze obserwacje przeprowadzone przez inne ośrodki badawcze). Autorzy sugerują, że jest to spowodowane przede wszystkim stresem i zmęczeniem związanymi z życiem w dużych miastach. Badanie to potwierdziło trend rosnącej potrzeby ludzi do odcinania się od cywilizacji i kontaktu z naturą, co jest ważne dla zrozumienia i rozwiązywania problemów związanych z życiem w ośrodkach zurbanizowanych.

W pracy naukowej *For the Love of Nature: Exploring the Importance of Species Diversity and Micro-Variables Associated with Favorite Outdoor Places* opublikowanym przez czasopismo akademickie *Frontiers in Psychology* [4], w dziale psychologii środowiskowej. Autorzy skupili się na zbadaniu znaczenia różnorodności gatunków i zmienności mikro-środowiskowych związanych z ulubionymi miejscami na świeżym powietrzu. Celem badania było zrozumienie, jak te czynniki wpływają na dobre samopoczucie i zdrowie ludzi. Artykuł zawiera wyniki badań przeprowadzonych za pomocą ankiet i wywiadów, które pozwoliły autorom na zebranie informacji o ulubionych miejscach na świeżym powietrzu i o czynnikach, które wpłynęły na ich wybór. Na Rysunku 1 zostały



Rysunek 1. Popularność różnych typów terenów zielonych wśród ulubionych miejsc uczestników badania [4]

przedstawione wyniki ankiety, podczasktórej respondentom zadano pytanie "Jakie są twoje ulubione miejsca na świeżym powietrzu w Australii Południowej?". Łącznie podano 1 022 ulubionych miejsc, przy czym respondenci zazwyczaj wymieniali od jednego do trzech ulubionych miejsc. Po pogrupowaniu ulubionych miejsc według nazwy, lista ta została zredukowana do 241 unikalnych lokalizacji. Każda unikalna lokalizacja została następnie sklasyfikowana przy użyciu zmodyfikowanej typologii parków. Niektóre odpowiedzi nie mogły być sklasyfikowane przy użyciu typologii, ponieważ były albo zbyt niejasne (np. "lokalny park"), albo odnosiły się do dużego regionu, który prawdopodobnie zawierał wiele typów terenów zielonych (np. "Wzgórza Adelajdy"). Te odpowiedzi zostały zakodowane jako "nieznane" lub "ogólny region", odpowiednio. Najczęściej wymienianymi ulubionymi miejscami były "parki przyrody", takie jak rezerваты i parki narodowe (39,52%). Badanie potwierdza że ludzie nie tylko uwielbiają kontakt z odseparowaną naturą (ponieważ tak też funkcjonują parki), ale też wskazują tę opcję najczęściej jako ich ulubione miejsce na świeżym powietrzu.

Pytanie pojawia się zatem następujące: czy możliwym byłoby poświęcenie życia w ośrodkach miejskich bez rezygnacji ze swojej pracy na rzecz przeprowadzki na tereny „bardziej dzikie”, które naturalnie wpływają korzystnie na jakość życia? Jeszcze 20 lat temu całkowite odcięcie się od dobrodziejstw jakie daje życie w dużych miastach stonowałoby barierę nie do przeskoczenia. Dostęp do edukacji, opieki medycznej, sklepów, rozrywki i możliwości zatrudnienia nieporównywalnie większe względem chociażby podmiejskich wsi to tylko niektóre z przykładów czynników, które brali pod uwagę młodzi dorośli wybierający miejsce zamieszkania, w którym będą chcieli przeżyć resztę życia. Jak jednak zweryfikował postęp technologiczny - w dobie wszechobecnego internetu, usług na rządanie „door-to-door” czy chociażby teleporad, wcześniej wymienione czynniki przemierzające „za” życiem w mieście tracą swoją pierwotną moc.

Kluczowym czynnikiem, który bez wątpienia wpłynął na taki stan rzeczy był niewątpliwie rozwój i rozpowszechnienie usług internetowych umożliwiające wielu pracownikom pracę zdalną. Dla wielu młodych, ale i nie tylko, komputer podłączony do stałego łącza sieciowego jest jedynym narzędziem pracy. Co więcej, powstały liczne badania potwierdzające pozytywny wpływ pracy zdalnej na satysfakcję z życia. Jak opublikowano w czasopiśmie *International Journal of Human Resource Management* w artykule pt. *Remote work in the COVID-19 era: exploring the work-life balance and job satisfaction of remote workers* [5] większość pracowników pracujących zdalnie w dobie pandemii COVID-19 odnotowała poprawę równowagi między życiem zawodowym a prywatnym. Jednocześnie, większość badanych wykazała również wyższe zadowolenie z pracy w porównaniu z okresem przed pandemią. Potwierdza to tezę mówiącą, że praca zdalna nie tylko jest możliwa, ale też, że wprost wpływa na poprawę ogólnej jakości życia.

„Choć ludzie, którzy w kilkaset tysięcy skupili się na niewielkim obszarze, starali się, jak tylko mogli, zeszpecić tę ziemię, na której się tłoczyli, choć wbijali w nią kamienie, żeby nic na niej nie rośło, choć wydzielali z niej każdą kiełkującą trawkę, choć dymili węglem kamiennym i naftą, choć obcinali drzewa i płoszyli wszystkie zwierzęta i ptaki - wiosna była wiosną nawet i w mieście. Słońce grzało, trawa powracając do życia rośla i zieleniła się wszędzie, gdzie jej nie powyrywano, nie tylko na trawnikach bulwarów, ale i między kamiennymi płytami; brzozy, topole, czeremchy rozwijały swe lepkie i wonne liście, pękały nabrzmiałe pączki lip; kawki, wróble i gołębie wiosennym obyczajem radośnie stały już sobie gniazda, a muchy brzęczały u ścian, przygrzane słońcem. Wesole były i rośliny, i ptaki, i owady, i dzieci. Ale ludzie dojrzały, dorośli ludzie nie przestawali oszukiwać i męczyć siebie i innych. Ludzie uważali, że święty i ważny jest nie ten wiosenny poranek, nie to piękno świata bożego, dane dla szczęścia wszystkich istot - piękno, które usposabia do pokoju, zgody i miłości, ale święte i ważne jest to, co oni sami wymyślili, żeby panować jedni nad drugimi.”[6]

Lew Tolstoj, *Zmartwychwstanie*, Rosja, 1899

*Wstawiona wyżej obserwacja Lwa Tolstoja zwraca uwagę na problem w społeczeństwie, który cechuje współczesnych w tamtym czasie dorosłych ludzi. Opisane zostały różnicę między naturalną radością i pięknem wiosny a ludzkim działaniem, które jest przeciwieństwem tej radości i piękna. Autor ukazuje, że wiosna jest wiosną nawet w mieście, mimo, że ludzie starają się z jakiegoś niesprecyzowanego powodu zniszczyć środowisko i zamienić je w chaotyczne i zanieczyszczone miejsce. Mimo tego, przyroda nadal kwitnie i jest radosna, zarówno wśród roślin, zwierząt, jak i dzieci. Ludzie natomiast, nawet dorośli, wciąż dążą do władzy i kontroli nad innymi, ignorując piękno i radość natury. Według autora, ludzie

nie doceniają tego, co naprawdę jest ważne i święte, takie jak piękno i radość wiosny, ale skupiają się na swoich własnych, często szkodliwych, wymyślonych celach. Jednak to jak etyczna jest taka działalność nie będzie poruszane w niniejszej pracy, a umieszczona została wyłącznie jako zwroć uwagi, że problem jest znacznie głębszy i istnieje nie tylko na przykładzie surowych badań, ale też w obserwacjach życia codziennego zwykłych ludzi bardziej wrażliwych na przyszłe losy świata.

Przedstawione wyżej argumenty pozwalają wyciągnąć wstępny bardzo ogólny wniosek, że w dzisiejszych czasach możliwe jest zrezygnowanie z życia w mieście na rzecz całkowitego odseparowania od cywilizacji. Powody takiego działania mogą być różne, od chęci poprawy jakości życia, co zostało wykazane we wspomnianych uprzednio badaniach, po osobiste powody etyczne, nie mniej jednak, jest to wykonalne, a na dodatek daje to wymierne korzyści. Należałoby tu oczywiście uwzględnić niezbędne czynniki jakimi byłyby uzyskanie bezprzewodowego dostępu do internetu możliwe dzięki chociażby technologii Starlink oraz dostęp do bezprzewodowej energii elektrycznej możliwy dzięki korzystaniu z odnawialnych źródeł energii jak energia słoneczna, wiatrowa czy geotermalna. Problem, który chciałbym poruszyć jest jednak inny, ponieważ skoro zostało już wstępnie ustalone, że takie życie byłoby możliwe to jakie byłoby idealne miejsce w Polsce by spełnić ten cel? Lub bardziej ogólnie - gdzie najlepiej byłoby szukać miejsca najbardziej dzikiego, nieskażonego ingerencją ludzką, najbardziej naturalnego i wolnego od cywilizacji? Część analityczna niniejszej pracy została poświęcona próbie identyfikacji polskich województw, które statystycznie najbardziej nadawałyby się do rozpoczęcia takich poszukiwań.



Rysunek 2. Wycinek z komiksu Tytus, Romek i A'Tomek autorstwa Henryka Jerzego Chmieleckiego

Opis zmiennych

W niniejszej analizie zostały wykorzystane dane pozyskane z internetowych baz udostępnionych przez Główny Urząd Statystyczny [7]. Siedemnaście zmiennych oznaczonych kolejno X1, X2 itd. Do badań zostały wybrane tylko takie dane, które jakkolwiek związane są z ingerencją, lub należą same w sobie do środowiska naturalnego (za wyjątkiem gęstości zaludnienia), które to jest głównym przedmiotem zainteresowania; dzięki czemu gdyby wszystkim destymulantom przypisać wartość równą zero, mielibyśmy do czynienia z terenem prawdopodobnie całkowicie dzikim. Omawiane zmienne to kolejno:

- X1 - Powierzchnia użytków rolnych w hektarach na 1000km,
- X2 - Liczba budynków mieszkalnych i mieszkalno-usługowych na 100 mieszkańców,
- X3 - Powierzchnia gruntów leśnych na 1000km²,*
- X4 - Gęstość zaludnienia powierzchni zabudowanej i zurbanizowanej, osoby na 10km²,
- X5 - Powierzchnia budynków niemieszkalnych, zbiorowego zamieszkania oraz domów letnich oddanych do użytkowania sumarycznie na 10 000 km²,
- X6 - Długość dróg publicznych ogółem na 100km²,
- X7 - Długość linii kolejowych normalnotorowych ogółem na 1000km²,
- X8 - Pobór wody na potrzeby gospodarki narodowej i ludności w ciągu roku w dekametrach na 100 mieszkańców,
- X9 - Ładunki zanieczyszczeń w ściekach odprowadzonych do wód lub do ziemi w kilogramach rocznie na 100 mieszkańców,
- X10 - Procentowy udział obszarów prawnie chronionych w powierzchni ogółem,*,
- X11 - Odpady wytworzone i dotychczas składowane z wyłączeniem odpadów komunalnych w tonach na 10 mieszkańców,
- X12 - Roczna emisja zanieczyszczeń do powietrza w tonach na 10 mieszkańców,
- X13 - Roczna emisja zanieczyszczeń pyłowych w kg na 100 mieszkańców,
- X14 - Powierzchnia dzikich wysypisk w m² na każdy km²,
- X15 - Pozyskanie drewna w m³ na hektar lasu,
- X16 - Hektary gruntów przeznaczonych pod budowę na 100km²,
- X17 - Liczba pojazdów mechanicznych na 10 mieszkańców.

*Zaznaczone kolorem zielonym zmienne to stymulanty, wszystkie pozostałe są destymulantami..

Zbiór danych

Wszystkie obliczenia do poniższej analizy znajdują się w drugim pliku dołączonym do tego sprawozdania.

Tabela 1. Zbiorcze zebranie ustandaryzowanych wartości wszystkich zmiennych biorących udział w badaniu. Opracowanie własne w programie Excel.

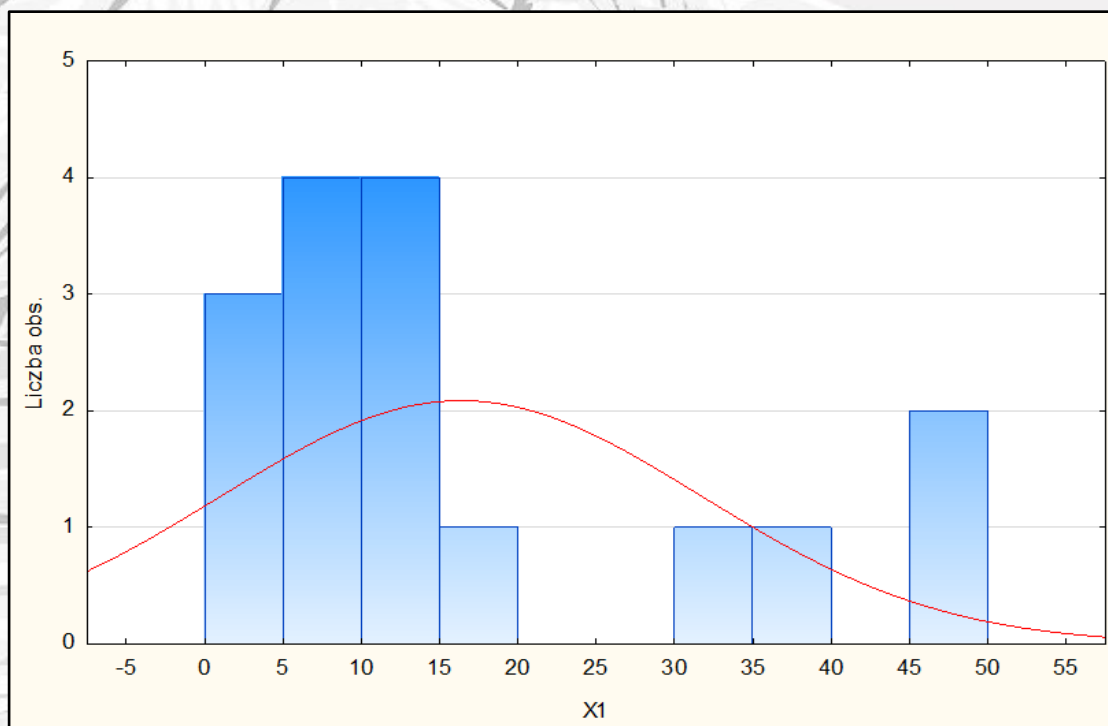
	D	D	S	D	D	D	D	D	S	D	D	D	D	D	D	D	D	D
WOJEWÓDZTWA	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	
POLSKA	17,572	17,643	59,879	12,100	91,047	137,500	61,808	24,110	23,704	32,300	28,023	54,502	57,673	6,649	45,952	8,740	18,374	
Dolnośląskie	15,367	14,297	14,735	14,500	97,244	127,500	89,036	12,861	22,871	18,600	120,566	52,013	48,275	22,079	157,678	16,577	17,933	
Kujawsko-pomorskie	4,908	15,344	29,425	11,200	77,171	151,900	66,715	12,236	67,735	32,200	7,336	47,815	89,061	8,258	94,119	3,239	18,080	
Lubelskie	5,663	18,771	75,173	13,100	113,356	144,800	59,279	11,642	66,242	19,500	30,805	156,722	64,645	6,018	46,526	10,599	18,712	
Lubuskie	11,153	22,281	96,837	8,100	53,625	154,400	43,508	14,409	24,610	22,700	25,421	23,027	55,101	2,444	32,698	3,274	19,051	
Łódzkie	36,066	15,724	10,927	7,000	51,559	113,000	65,270	8,766	3,992	37,400	6,001	22,997	60,015	4,914	129,474	12,328	19,281	
Małopolskie	5,219	20,800	132,386	22,600	121,535	207,500	71,132	12,981	8,879	53,100	12,749	22,674	34,246	0,434	43,423	23,593	17,526	
Mazowieckie	13,345	18,206	105,867	15,500	143,532	157,000	48,259	53,323	10,952	29,700	11,757	62,525	37,240	1,658	34,847	3,593	20,997	
Opolskie	2,836	18,022	15,399	10,100	77,359	112,300	83,298	13,459	28,756	27,600	11,865	176,665	120,782	8,962	141,711	9,272	18,583	
Podkarpackie	7,458	22,146	83,847	11,700	55,126	121,800	54,802	8,065	6,916	44,900	3,395	9,985	42,488	1,374	43,328	6,780	17,458	
Podlaskie	30,089	20,314	102,868	5,700	44,659	135,700	37,598	8,460	7,128	31,600	7,495	18,477	41,302	10,061	19,774	3,293	17,343	
Pomorskie	12,622	14,526	44,180	12,900	92,921	124,300	66,193	9,271	5,329	33,000	7,909	29,285	30,290	3,278	36,768	15,255	17,885	
Śląskie	3,002	15,084	67,733	35,500	204,994	201,700	151,788	9,446	37,670	22,100	60,752	78,311	93,352	45,978	57,048	26,941	15,978	
Świętokrzyskie	7,318	22,888	82,592	10,100	48,458	149,200	61,651	100,030	77,840	64,900	40,506	116,933	104,202	6,351	31,134	2,279	18,405	
Warmińsko-mazurskie	47,325	13,840	27,704	5,700	42,504	91,400	45,133	8,393	5,211	46,700	7,853	11,960	38,759	0,495	84,685	5,540	16,555	
Wielkopolskie	10,040	18,495	30,079	11,700	148,118	141,700	63,066	36,410	6,245	29,600	9,273	27,240	41,809	0,903	66,160	4,432	20,537	
Zachodniopomorskie	47,776	12,826	11,986	7,200	52,573	85,600	51,110	56,397	17,285	21,800	29,298	38,641	86,175	5,787	119,190	10,094	16,977	

Tabela 2. Podstawowe statystyki zmiennych biorących udział w badaniu. Opracowanie własne w programie Excel

Zmienna	Średnia	Mediana	Minimum	Maksimum	Odychlenie Standardowe	Współczynnik Zmienności
X1	16,26	10,60	2,84	47,78	14,80	91,03
X2	17,72	18,11	12,83	22,89	3,19	18,00
X3	58,23	55,96	10,93	132,39	38,51	66,14
X4	12,66	11,45	5,70	35,50	7,20	56,84
X5	89,05	77,26	42,50	204,99	45,30	50,87
X6	138,74	138,70	85,60	207,50	32,26	23,26
X7	66,11	62,36	37,60	151,79	25,86	39,11
X8	23,51	12,55	8,07	100,03	24,97	106,23
X9	24,85	14,12	3,99	77,84	24,02	96,65
X10	33,46	30,65	18,60	64,90	12,65	37,81
X11	24,56	11,81	3,39	120,57	29,02	118,17
X12	55,95	33,96	9,98	176,66	49,77	88,95
X13	61,73	51,69	30,29	120,78	27,30	44,23
X14	8,06	5,35	0,43	45,98	11,09	137,56
X15	71,16	51,79	19,77	157,68	42,92	60,32
X16	9,82	8,03	2,28	26,94	7,27	74,08
X17	18,21	18,01	15,98	21,00	1,29	7,09

Ostatnia kolumna tabeli 2, zwana Współczynnikiem zmienności, jest istotnym elementem analizy danych. Ten współczynnik jest wynikiem podzielenia odchylenia standardowego przez średnią arytmetyczną.

Z punktu widzenia badania, wartości zmienności zawierające się w zakresie od 10% do 100% są uważane za odpowiednie. Wartość niższa niż 10% oznacza niską zmienność i jednorodność badanej populacji, co może prowadzić do nieznaczących różnic wyników analizy. Natomiast wartość powyżej 100% wskazuje na dużą niejednorodność badanej populacji, co może powodować zniekształcenie wyników i błędne wnioski. Wynika z tego, że badanie na danych wybranych w tej analizie wskazują na konieczność wykluczenia zmiennej X17 czyli liczby pojazdów mechanicznych na 10 mieszkańców, z racji niskiego współczynnika zmienności, niższego niż 10%. Nie będzie on uwzględniany w dalszej części analizy.



Rysunek 3. Histogram asymetrii zmiennej X1. Opracowanie własne w programie Statistica

W badaniach brana jest pod uwagę wpływ poszczególnych zmiennych na środowisko naturalne, uwzględniając czy wpływają one na nie stymulująco czy destymulująco. Większość zmiennych ma negatywny wpływ na środowisko. Jednak dwie zmienne, tj. Powierzchnia gruntów leśnych (X3) i procentowy udział obszarów prawnie chronionych (X10), są uważane za stymulujące ze względu na ich pozytywny wpływ na dzikość natury, zarówno obiektywnie, jak i subiektywnie.

Metody porządkowania liniowego służą do określenia pozycji lub klasyfikacji obiektów w zależności od jakiegoś kryterium. W tym badaniu badane są województwa, a kryteriami są powyższe zmienne. W badaniu wykorzystano następujące metody porządkowania: Metoda Wzorca, Metoda Wzorca z wagami, Metoda TOPSIS i Metoda Sumy Rang. Metoda Wzorca i TOPSIS to metody wzorcowe, gdzie obiekty są porównywane z punktem odniesienia, czyli wzorcem. W metodzie TOPSIS ustala się dwa punkty odniesienia: wzorzec i antywzorzec. Wzorzec jest reprezentowany przez wartość maksymalną stymulant i minimalną destymulant, a antywzorzec jest reprezentowany przez wartość minimalną stymulant i maksymalną destymulant.

Metoda Hellwiga i wzorca z wagami polegają na wyznaczeniu sumy odległości między wzorcami i antywzorcami. Druga metoda, czyli wzorca z wagami, uwzględnia dodatkowo stosunek współczynnika zmienności danej zmiennej do sumy wszystkich współczynników zmienności [8].

Tabele 3 – 6. Wyniki analiz porządkowania liniowego metodami Wzorca, Wzorca z wagami, TOPSIS oraz Sumy Rang. Opracowanie własne w programie Excel

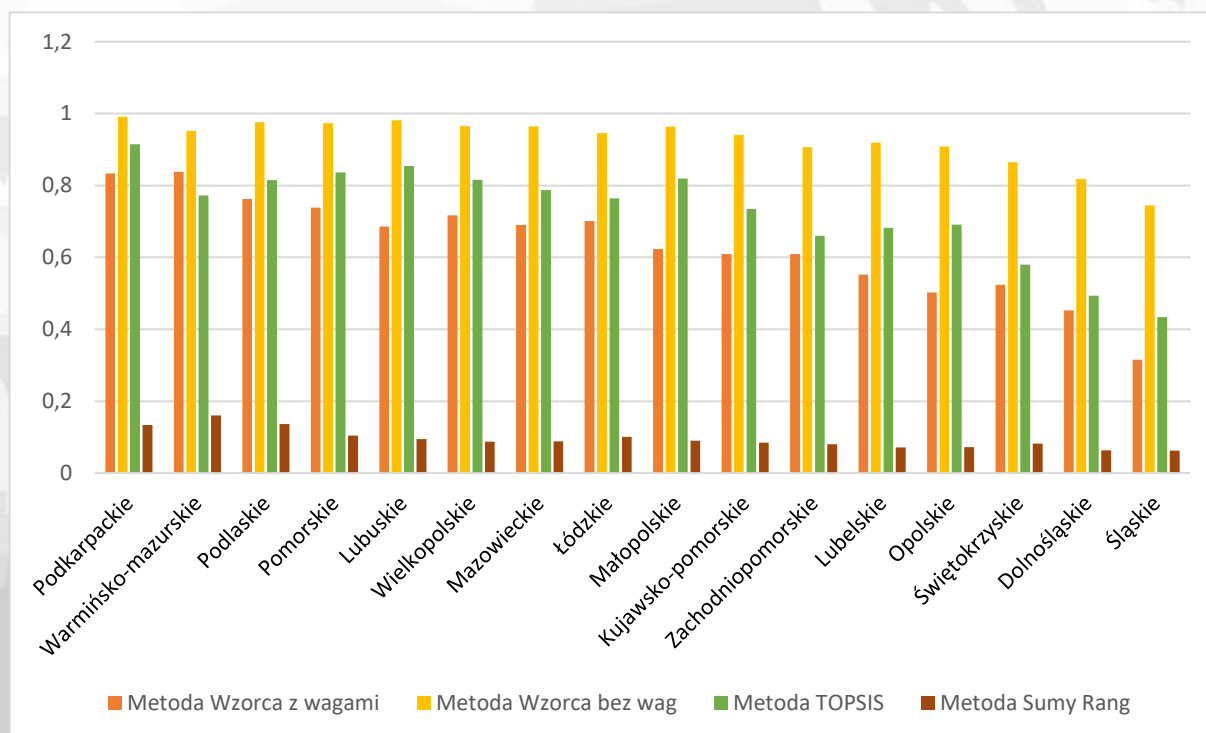
Metoda Wzorca bez wag	
Województwo	mi
Warmińsko-mazurskie	0,8383478
Podkarpackie	0,8332504
Podlaskie	0,762699
Pomorskie	0,7380101
Wielkopolskie	0,7173453
Łódzkie	0,7006345
Mazowieckie	0,6903643
Lubuskie	0,6857163
Małopolskie	0,6238114
Kujawsko-pomorskie	0,6096294
Zachodniopomorskie	0,6095256
Lubelskie	0,5523438
Świętokrzyskie	0,523497
Opolskie	0,5023395
Dolnośląskie	0,4520778
Śląskie	0,3152595

Metoda Wzorca z wagami	
Województwa	mi
Podkarpackie	0,9912878
Lubuskie	0,981177
Podlaskie	0,9763726
Pomorskie	0,9736955
Wielkopolskie	0,9656999
Mazowieckie	0,9644396
Małopolskie	0,9637532
Warmińsko-mazurskie	0,9518388
Łódzkie	0,9459693
Kujawsko-pomorskie	0,9402548
Lubelskie	0,9187949
Opolskie	0,9082991
Zachodniopomorskie	0,9064173
Świętokrzyskie	0,8652171
Dolnośląskie	0,8187597
Śląskie	0,7447042

Metoda TOPSIS	
Województwa	qi
Podkarpackie	0,9143383
Lubuskie	0,8544337
Pomorskie	0,8366529
Małopolskie	0,8196322
Wielkopolskie	0,8159008
Podlaskie	0,8151297
Mazowieckie	0,7876447
Warmińsko-mazurskie	0,7720354
Łódzkie	0,764588
Kujawsko-pomorskie	0,7350776
Opolskie	0,6915275
Lubelskie	0,6821087
Zachodniopomorskie	0,6600942
Świętokrzyskie	0,579991
Dolnośląskie	0,4931246
Śląskie	0,4336468

Metoda Sumy Rang	
Województwo	Suma Rang
Warmińsko-mazurskie	75
Podlaskie	88
Podkarpackie	90
Pomorskie	115
Łódzkie	119
Lubuskie	127
Małopolskie	133
Mazowieckie	136
Wielkopolskie	137
Kujawsko-pomorskie	142
Świętokrzyskie	147
Zachodniopomorskie	150
Opolskie	166
Lubelskie	168
Dolnośląskie	190
Śląskie	193

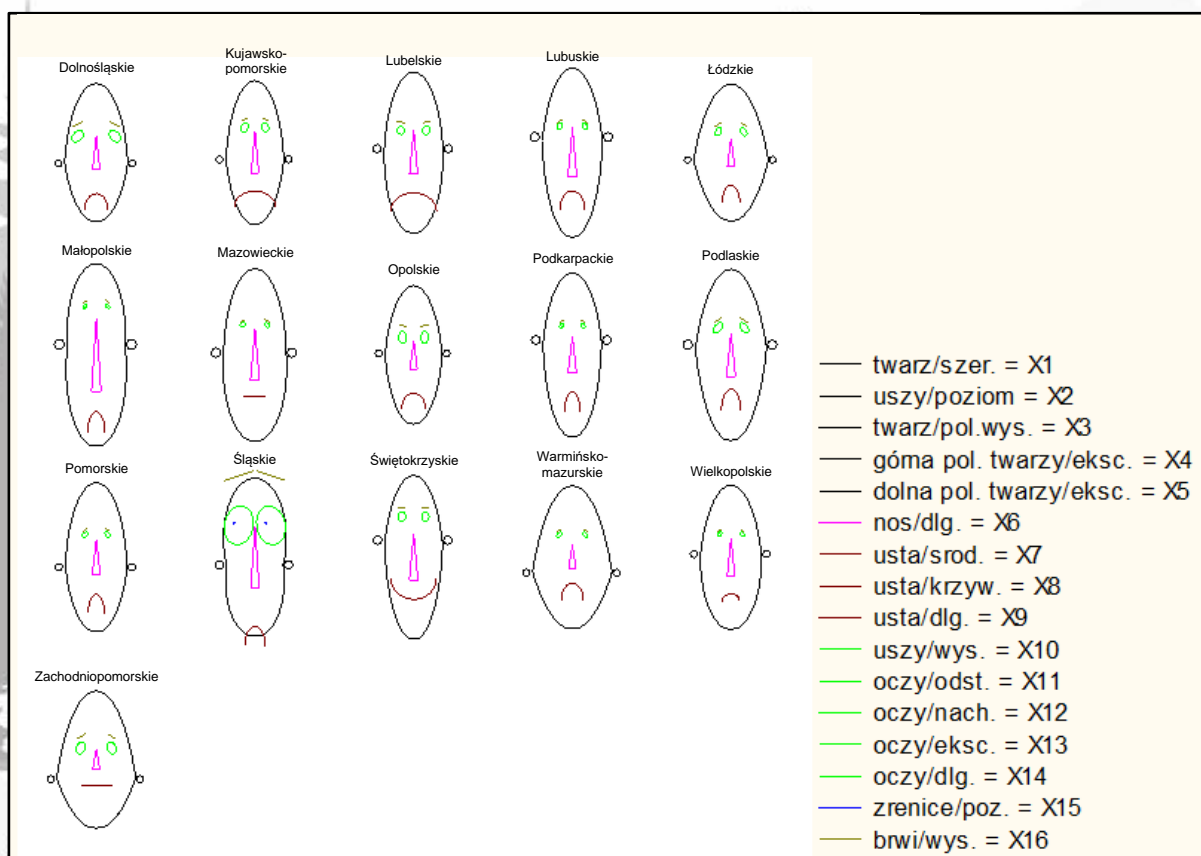
Z powyższych wyników analizy z wykorzystaniem porządkowania liniowego metodami Wzorca (zarówno z wagami jak i bez), Topsis oraz Sumy rang można wywnioskować że województwa: Podkarpackie, Lubuskie, Podlaskie, Pomorskie oraz Warmińsko-mazurskie są warte bliższej obserwacji w kontekście wytypowania regionu najmniej zniszczonego ingerencją człowieka w środowisko naturalne. Poniżej na wykresie zostały przedstawione wyniki analizy w formie wykresu słupkowego.



Rysunek 4. Wykres wyników porządkowania liniowego poszczególnych województw. Metoda Sumy Rang została odpowiednio przekształcona dla lepszej wizualizacji. Opracowanie własne w programie Excel.

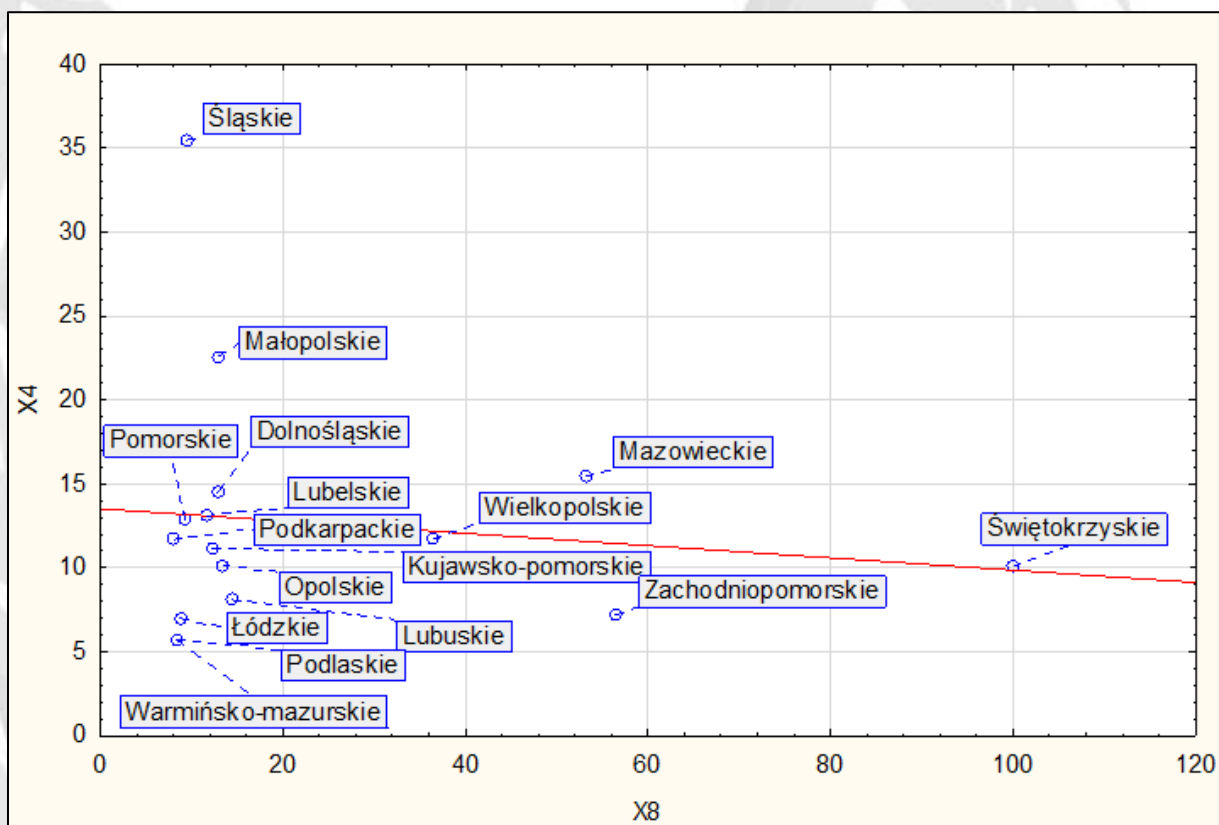
W badaniu również warto uwzględnić wykres Chernoffa. Jest to technika wizualizacji danych, która wykorzystuje rysunki ludzkich twarzy do przedstawienia informacji. Każda cecha twarzy jest używana do reprezentowania różnych wartości danych, takich jak średnia, odchylenie standardowe i skośność. Charakterystyka wykresu Chernoffa jest następująca:

- ❖ Reprezentacja wielu wymiarów danych: Każdy element twarzy, taki jak oczy, nos, usta i brwi, jest używany do reprezentowania różnych wartości danych.
- ❖ Możliwość wizualnego porównania: Pozwala porównywać wiele obiektów na raz, co jest trudne do osiągnięcia w przypadku innych wizualizacji danych.
- ❖ Wykazanie tendencji i odchyłeń: Wykresy Chernoffa umożliwiają wykazanie tendencji i odchyłeń w danych.
- ❖ Zwięzła i przejrzysta prezentacja: Dzięki prostocie rysunków twarzy, wykres Chernoffa jest łatwy do zrozumienia i przedstawia informacje w skondensowanej formie.



Rysunek 5. Wykres Twarzy Chernoffa. Opracowanie własne w programie Statistica.

Na rysunku 5 został przedstawiony wykres Chernoffa dla danych biorących udział w analizie. Pomimo faktu, że na pierwszy rzut oka ciężko wychwycić duże różnice to warto zwrócić uwagę na rysunek znajdujący się w trzecim rzędzie, drugiej kolumny. Twarz z bardzo wytrzeszczonymi, zdziwionymi oczami oraz wyjątkowo smutną miną reprezentuje województwo Śląskie, które zdecydowanie najsłabiej wypadło w każdej wcześniejszej analizie przy wykorzystaniu którejkolwiek z metod. Wskazuje to, że województwo śląskie jest województwem prawdopodobnie najbardziej naruszone przez człowieka przez ingerencję w środowisko naturalne



Rysunek 6. Wykres rozrzutu zmiennej X4 względem zmiennej X8. Opracowanie własne w programie Statistica.

Na rysunku 6 przedstawiono wykres rozrzutu zmiennej X4 czyli gęstości zaludnienia powierzchni zabudowanej i zurbanizowanej względem zmiennej X8 wskazującej na pobór wody na potrzeby gospodarki narodowej i ludności w ciągu roku w dekametrach na 100 mieszkańców. Większość województw jest skupiona w jednym miejscu wykazując względnie podobne relacje zagęszczenia ludności do poboru wody. Warto tu jednak zwrócić uwagę na województwa wyraźnie oddalone od reszty, czyli Śląskie i Świętokrzyskie. Województwo śląskie znajdujące się w górnej części wykresu wskazuje na wyjątkowo mały pobór wody pomimo największej spośród wszystkich województw gęstości zaludnienia. Województwo świętokrzyskie natomiast wykazuje największy pobór wody, pomimo względnie średniej na tle pozostałych jednostek administracyjnych gęstości zaludnienia. Pobór wody w tym województwie jest prawie dwukrotnie większy niż w Mazowieckim i Zachodniopomorskim oraz ponad dziesięciokrotnie większy niż w Śląskim, które ma największą gęstość zaludnienia.

Analiza głównych składowych PCA

Analiza głównych składowych (PCA) jest to powszechnie używana technika statystyczna, która polega na redukcji wymiarów danych. Jest to szczególnie przydatne, gdy dane mają wiele cech i trudno je zinterpretować i zrozumieć. PCA polega na znalezieniu tzw. głównych składowych danych, czyli nowych zmiennych, które opierają się na kombinacjach oryginalnych cech i które kumulują jak największą wariancję w danych. W efekcie, dane są zredukowane do znacznie mniejszej liczby zmiennych, które umożliwiają łatwiejsze zrozumienie i interpretację danych.

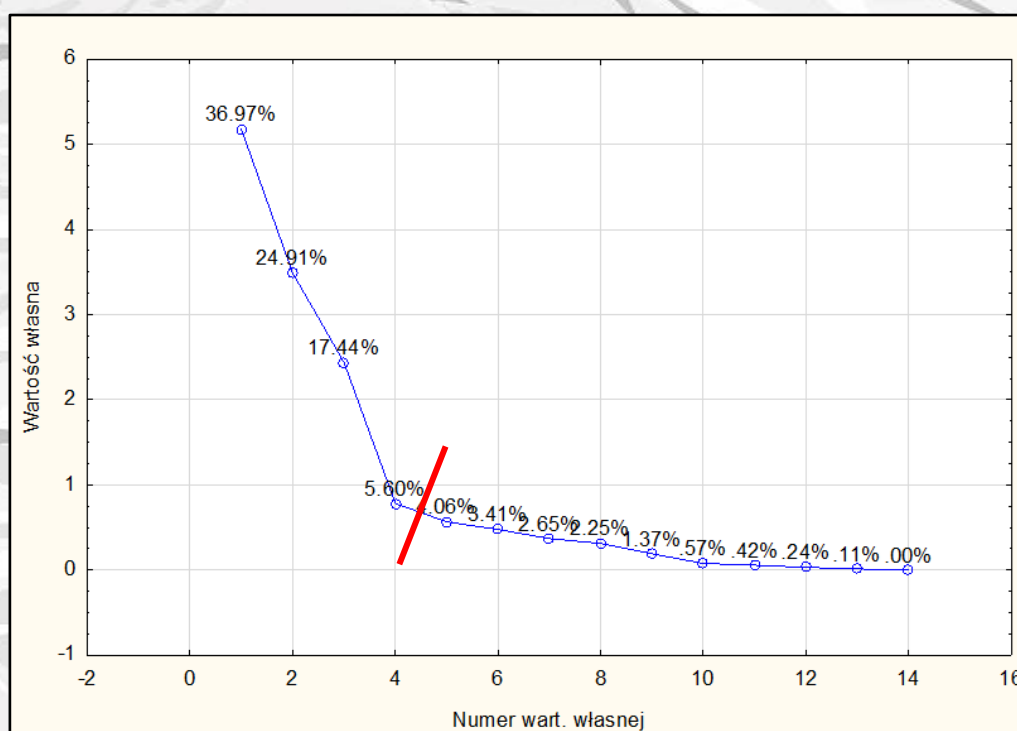
Tabela 7. Macierz korelacji zmiennych. Opracowanie własne w programie Statistica.

Zmienne	Średnia	Odch. Stand	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16
X1	16,2615	15,28748	1,000000	-0,484049	-0,379549	-0,535707	-0,514402	-0,678320	-0,422394	0,025434	-0,450668	-0,002053	-0,127487	-0,451383	-0,210973	-0,221443	0,286166	-0,183871
X2	17,7227	3,29488	-0,484049	1,000000	0,733540	-0,065797	-0,141081	0,392514	-0,300952	0,232235	0,171306	0,436006	-0,239337	0,104456	-0,049408	-0,316651	-0,619738	-0,331511
X3	58,2336	39,77644	-0,379549	0,733540	1,000000	0,304677	0,177733	0,676063	-0,164142	0,093793	0,028732	0,319494	-0,180099	-0,089667	-0,330331	-0,114981	-0,832501	0,019801
X4	12,6625	7,43316	-0,535707	-0,065797	0,304677	1,000000	0,859851	0,781081	0,842831	-0,127038	0,127167	-0,115460	0,362147	0,147588	0,093769	0,708951	-0,160761	0,778672
X5	89,0458	46,78541	-0,514402	-0,141081	0,177733	0,859851	1,000000	0,671035	0,690650	-0,071859	0,039934	-0,330495	0,263231	0,194082	-0,030383	0,543660	-0,137792	0,565515
X6	138,7375	33,32292	-0,678320	0,392514	0,676063	0,781081	0,671035	1,000000	0,475472	-0,026629	0,257475	0,098940	0,140369	0,068555	-0,029984	0,377955	-0,485085	0,449322
X7	66,1149	26,70727	-0,422394	-0,300952	-0,164142	0,842831	0,690650	0,475472	1,000000	-0,180843	0,211936	-0,233463	0,514234	0,304879	0,411338	0,881647	0,251411	0,759666
X8	23,5093	25,79359	0,025434	0,232235	0,093793	-0,127038	-0,071859	-0,026629	-0,180843	1,000000	0,372413	0,377408	0,084227	0,240703	0,330294	-0,163577	-0,160802	-0,363155
X9	24,8538	24,81015	-0,450668	0,171306	0,028732	0,127167	0,039934	0,257475	0,211936	0,372413	1,000000	0,059803	0,262321	0,653062	0,665167	0,246271	-0,057389	-0,137147
X10	33,4625	13,06654	-0,002053	0,436006	0,319494	-0,115460	-0,330495	0,098940	-0,233463	0,377408	0,059803	1,000000	-0,342779	-0,153811	-0,055767	-0,388208	-0,324037	-0,150357
X11	24,5613	29,97606	-0,127487	-0,239337	-0,180099	0,362147	0,263231	0,140369	0,514234	0,084227	0,262321	-0,342779	1,000000	0,202527	0,163691	0,640921	0,376570	0,408193
X12	55,9542	51,40368	-0,451383	0,104456	-0,089667	0,147588	0,194082	0,068555	0,304879	0,240703	0,653062	-0,153811	0,202527	1,000000	0,700273	0,229597	0,164368	0,029979
X13	61,7338	28,19877	-0,210973	-0,049408	-0,330331	0,093769	-0,030383	-0,029984	0,411338	0,330294	0,665167	-0,055767	0,163691	0,700273	1,000000	0,391859	0,365183	-0,010487
X14	8,0621	11,45420	-0,221443	-0,316651	-0,114981	0,708951	0,543660	0,377955	0,881647	-0,163577	0,246271	-0,388208	0,640921	0,229597	0,391859	1,000000	0,185550	0,595654
X15	71,1602	44,33087	0,286166	-0,619738	-0,832501	-0,160761	-0,137792	-0,485085	0,251411	-0,160802	-0,057389	-0,324037	0,376570	0,164368	0,365183	0,185550	1,000000	0,178863
X16	9,8180	7,51122	-0,183871	-0,331511	0,019801	0,778672	0,565515	0,449322	0,759666	-0,363155	-0,137147	-0,150357	0,408193	0,029979	-0,010487	0,595654	0,178863	1,000000

Pierwszym krokiem analizy było utworzenie macierzy korelacji zmiennych w celu wytypowania i odrzucenia zmiennych na podstawie niewystarczającego skorelowania ze sobą cech. Macierz korelacji dla zmiennych biorących udział w tej analizie została przedstawiona w tabeli 7. Na żółto zaznaczone zostały zmienne X8 i X10, które ze względu na brak wykazanej korelacji względem innych zmiennych zostają odrzucone i nie będą brać udziału w dalszej części badania.

Tabela 8. Wartości własne wraz z podstawowymi statystykami. Opracowanie własne w programie Statistica.

Nr wartości	Wartość wł	% ogółu varianc.	Skumul wartość wł	Skumul. %
1	5,176285	36,97347	5,17629	36,9735
2	3,488042	24,91458	8,66433	61,8880
3	2,440984	17,43560	11,10531	79,3236
4	0,784361	5,60258	11,88967	84,9262
5	0,568179	4,05842	12,45785	88,9847
6	0,477577	3,41127	12,93543	92,3959
7	0,370941	2,64958	13,30637	95,0455
8	0,314616	2,24726	13,62099	97,2928
9	0,192044	1,37174	13,81303	98,6645
10	0,079125	0,56518	13,89215	99,2297
11	0,058887	0,42062	13,95104	99,6503
12	0,033463	0,23902	13,98450	99,8893
13	0,015428	0,11020	13,99993	99,9995
14	0,000067	0,00048	14,00000	100,0000



Rysunek 7. Wykres osypiska wartości własnych. Opracowanie własne w programie Statistica.

Po przeprowadzeniu analizy skumulowanej wartości wariancji za pomocą metody Kaisera, okazało się, że cztery główne składowe opisują 84,92% zmienności, dlatego te cztery składowe zostały wybrane do dalszej części badania. Wykres osypiska po czwartej składowej bardzo się spłaszcza co też wskazuje na brak istotności kolejnych wartości.

Tabela 9. Wektory własne macierzy korelacji, tylko pierwsze cztery czynniki.
Opracowanie własne w programie Statistica.

Zmienne	Czynn. 1	Czynn. 2	Czynn. 3	Czynn. 4
X1	0,274971	0,245179	0,229860	-0,165975
X2	0,030633	-0,432504	-0,239439	-0,189379
X3	-0,066641	-0,493734	0,032900	-0,210385
X4	-0,408714	-0,085229	0,174352	0,105018
X5	-0,352549	-0,075946	0,178413	0,324130
X6	-0,317604	-0,325122	0,052340	-0,024217
X7	-0,402699	0,158331	0,063551	0,097051
X9	-0,161977	-0,003992	-0,515574	-0,187580
X11	-0,235928	0,211572	0,008766	-0,770168
X12	-0,177038	0,079882	-0,483431	0,216253
X13	-0,150680	0,213917	-0,473169	0,145881
X14	-0,358860	0,174903	0,053775	-0,255513
X15	-0,001079	0,478720	-0,020894	0,038708
X16	-0,316082	0,092551	0,304838	0,055266

Tabela 10. Korelacja czynników zmiennych, tylko pierwsze cztery czynniki.
Opracowanie własne w programie Statistica.

Zmienne	Czynn. 1	Czynn. 2	Czynn. 3	Czynn. 4
X1	0,625599	0,457904	0,359125	-0,146995
X2	0,069695	-0,807757	-0,374092	-0,167722
X3	-0,151618	-0,922112	0,051401	-0,186326
X4	-0,929884	-0,159176	0,272401	0,093009
X5	-0,802099	-0,141839	0,278747	0,287063
X6	-0,722594	-0,607208	0,081773	-0,021448
X7	-0,916199	0,295704	0,099289	0,085953
X9	-0,368521	-0,007456	-0,805515	-0,166129
X11	-0,536771	0,395137	0,013696	-0,682093
X12	-0,402787	0,149189	-0,755295	0,191523
X13	-0,342820	0,399517	-0,739263	0,129198
X14	-0,816459	0,326654	0,084016	-0,226293
X15	-0,002456	0,894072	-0,032644	0,034281
X16	-0,719133	0,172850	0,476269	0,048946

W tabelach 9 i 10 zostały przedstawione macierze korelacji składowych i zmiennych dla czterech czynników które łącznie opisują prawie 85% zmienności. Można z nich wywnioskować że największy wpływ na pierwszą składową mają zmienne X4 – gęstość zaludnienia, X5 – powierzchnie budynków, X7 – długości linii kolejowych,

X14 – powierzchnie dzikich wysypisk, na drugą: X2 – liczba budynków, X3 – powierzchnie lasów, X15 – wycinka drewna, na trzecią X9 – ładunki zanieczyszczeń w ściekach, X12 – roczna emisja zanieczyszczeń powietrza, X13 – roczna emisja zanieczyszczeń pyłowych, a na czwartą X5 – powierzchnia budynków niemieszkalnych, X12 oraz X14. Warto tu zwrócić uwagę na drugi czynnik, który uwzględnia największy wpływ zmiennych odnoszących się mniej lub bardziej bezpośrednio do kwestii wykorzystania drewna lub ogólnej do pewnego rodzaju „drzewności” województwa. Jeszcze w przypadku pierwszego czynnika dałoby się skategoryzować zbiór tych zmiennych jako „przemysłowy”, lecz w pozostałych przypadkach ciężko jednoznacznie określić zbiory zmiennych.

Tabela 11. Wkład procentowy zmiennych na składowe. Opracowanie własne w programie Statistica.

Zmienne	Czynn. 1	Czynn. 2	Czynn. 3	Czynn. 4	Czynn. 5	Czynn. 6	Czynn. 7	Czynn. 8	Czynn. 9	Czynn. 10	Czynn. 11	Czynn. 12	Czynn. 13	Czynn. 14
X1	39%	60%	73%	75%	90%	90%	95%	96%	99%	99%	100%	100%	100%	100%
X2	0%	66%	80%	83%	84%	94%	95%	98%	98%	99%	100%	100%	100%	100%
X3	2%	87%	88%	91%	92%	93%	97%	97%	98%	99%	100%	100%	100%	100%
X4	86%	89%	96%	97%	98%	98%	98%	98%	98%	98%	99%	100%	100%	100%
X5	64%	66%	74%	82%	84%	95%	95%	97%	99%	100%	100%	100%	100%	100%
X6	52%	89%	90%	90%	90%	90%	91%	94%	98%	99%	99%	100%	100%	100%
X7	84%	93%	94%	94%	95%	96%	98%	98%	99%	99%	99%	100%	100%	100%
X9	14%	14%	78%	81%	84%	90%	90%	99%	99%	99%	100%	100%	100%	100%
X11	29%	44%	44%	91%	98%	99%	99%	99%	99%	100%	100%	100%	100%	100%
X12	16%	18%	75%	79%	82%	82%	97%	99%	99%	100%	100%	100%	100%	100%
X13	12%	28%	82%	84%	91%	96%	97%	97%	98%	99%	100%	100%	100%	100%
X14	67%	77%	78%	83%	91%	91%	92%	97%	98%	99%	100%	100%	100%	100%
X15	0%	80%	80%	80%	89%	92%	93%	94%	99%	99%	99%	100%	100%	100%
X16	52%	55%	77%	78%	78%	87%	93%	98%	99%	100%	100%	100%	100%	100%

Tabela 12. Zasoby zmiennej wspólnej, wartości wyrażone w %. Opracowanie własne w programie Statistica.

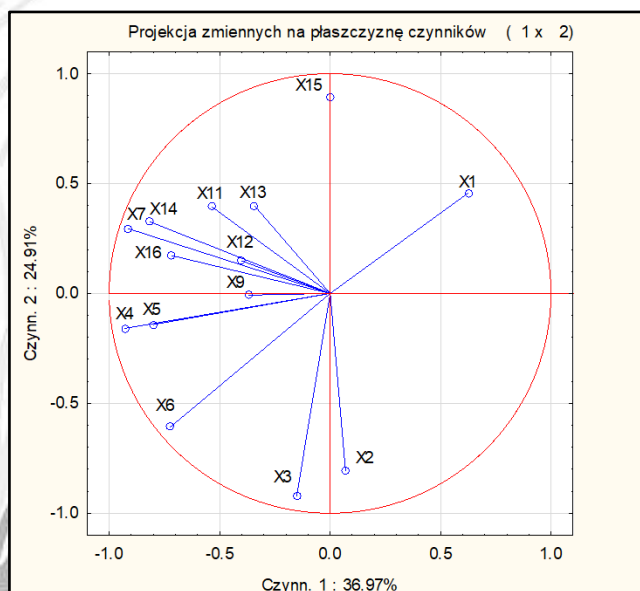
Zmienne	Czynn. 1	Czynn. 2	Czynn. 3	Czynn. 4	Czynn. 5	Czynn. 6	Czynn. 7	Czynn. 8	Czynn. 9	Czynn. 10	Czynn. 11	Czynn. 12	Czynn. 13	Czynn. 14
X1	8%	6%	5%	3%	26%	0%	15%	0%	16%	8%	8%	2%	2%	0%
X2	0%	19%	6%	4%	3%	21%	3%	7%	1%	17%	12%	5%	3%	1%
X3	0%	24%	0%	4%	2%	2%	10%	0%	5%	13%	13%	4%	0%	21%
X4	17%	1%	3%	1%	1%	0%	0%	0%	0%	3%	16%	23%	3%	31%
X5	12%	1%	3%	11%	3%	22%	0%	7%	12%	9%	1%	0%	13%	7%
X6	10%	11%	0%	0%	0%	0%	3%	11%	20%	5%	11%	15%	9%	5%
X7	16%	3%	0%	1%	1%	3%	4%	1%	4%	2%	2%	3%	32%	27%
X9	3%	0%	27%	4%	5%	14%	0%	28%	1%	1%	3%	11%	3%	2%
X11	6%	4%	0%	59%	12%	2%	1%	1%	0%	4%	4%	5%	1%	0%
X12	3%	1%	23%	5%	5%	0%	42%	6%	0%	4%	3%	0%	5%	3%
X13	2%	5%	22%	2%	13%	10%	2%	0%	7%	4%	17%	13%	2%	0%
X14	13%	3%	0%	7%	14%	0%	3%	18%	2%	20%	6%	0%	12%	3%
X15	0%	23%	0%	0%	15%	8%	2%	3%	23%	8%	0%	15%	2%	1%
X16	10%	1%	9%	0%	0%	20%	15%	16%	7%	2%	2%	3%	13%	0%

Biorąc pod uwagę wkład procentowy zmiennych na pierwszą składową, zmienne X4, X5, X7 i X14 odpowiadają za 58% całkowitej zmienności. W przypadku drugiej składowej wymienione wcześniej zmienne odpowiadają za 66%, dla trzeciej jest to 72%, a dla czwartej 70%. Najmniej istotną zmienną okazały się X1 – Powierzchnia użytków rolnych, X6 – długość dróg publicznych, X11 – odpady wytworzone i składowane oraz X16 – grunty przeznaczone pod budowę. Ciekawy jest tu przykład zmiennych X6 i X7 gdzie, jakby się mogło wydawać podobne zmienne, czyli długości odpowiednio dróg publicznych i linii kolejowych nie wykazują podobnej istotności.

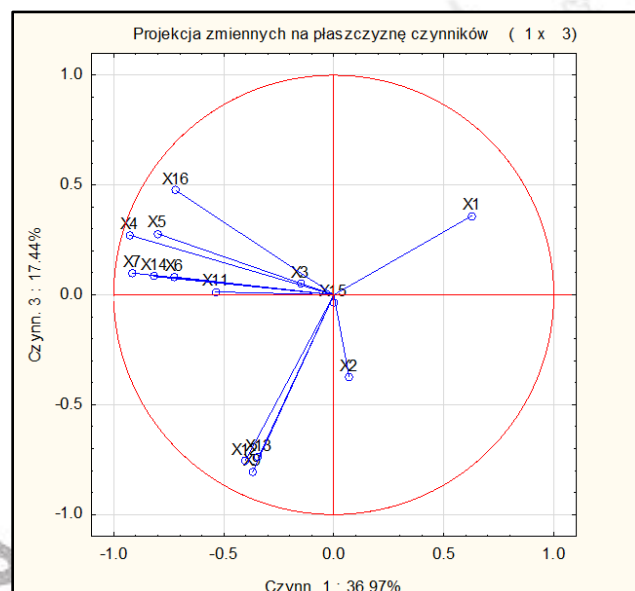
Równania głównych składowych są wyznaczane w celu uzyskania nowych reprezentacji danych, które są bardziej zwarte i łatwiejsze do interpretacji. W wielu przypadkach, dane mają wiele zmiennych, co może być trudne do analizy i interpretacji. PCA pozwala na redukcję liczby zmiennych, zachowując jednocześnie jak najwięcej informacji. Redukcja wymiarów jest szczególnie ważna w przypadku, gdy dane mają wiele zmiennych, a każda zmienna jest skorelowana z kilkoma innymi zmiennymi. W takim przypadku, każda zmienna jest do pewnego stopnia redundantą i nie wnoszą one wiele dodatkowej informacji do danych. PCA pozwala na usunięcie tych redundancji i uzyskanie nowych reprezentacji danych, które są bardziej zwarte i łatwiejsze do interpretacji. Oprócz redukcji wymiarów, równania głównych składowych są także wykorzystywane w wielu innych zastosowaniach, takich jak klasyfikacja, grupowanie danych, wizualizacja danych i wiele innych. Są one szczególnie przydatne w analizie danych, gdy celem jest zrozumienie zależności między różnymi zmiennymi i wykrycie wzorców w danych. Równania dla 4 najbardziej istotnych składowych zostały przedstawione poniżej.

Równania 1 - 4. Równania czterech pierwszych głównych składowych. Opracowanie własne.

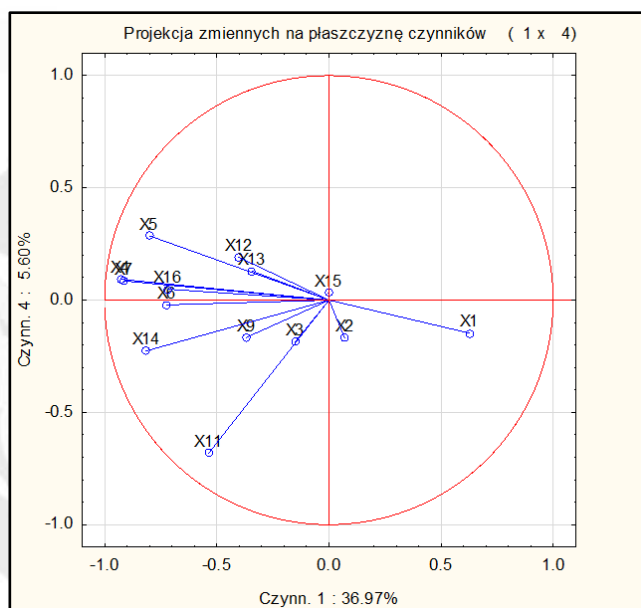
$$\begin{aligned}
 Y_1 &= -0,28X_1 - 0,03X_2 + 0,07X_3 + 0,41X_4 + 0,35X_5 + 0,32X_6 + 0,40X_7 + 0,16X_9 + 0,24X_{11} + 0,18X_{12} + 0,15X_{13} + 0,36X_{14} + 0,001X_{15} + 0,32X_{16} \\
 &\quad -0,63 \quad -0,07 \quad 0,15 \quad 0,93 \quad 0,80 \quad 0,72 \quad 0,92 \quad 0,37 \quad 0,54 \quad 0,40 \quad 0,34 \quad 0,82 \quad 0,002 \quad 0,72 \\
 Y_2 &= 0,25X_1 - 0,43X_2 - 0,49X_3 - 0,09X_4 - 0,08X_5 - 0,33X_6 + 0,16X_7 - 0,004X_9 + 0,21X_{11} + 0,08X_{12} + 0,21X_{13} + 0,17X_{14} + 0,48X_{15} + 0,09X_{16} \\
 &\quad 0,46 \quad -0,81 \quad -0,92 \quad -0,16 \quad -0,14 \quad -0,61 \quad 0,30 \quad -0,01 \quad 0,40 \quad 0,15 \quad 0,40 \quad 0,33 \quad 0,89 \quad 0,17 \\
 Y_3 &= 0,23X_1 - 0,24X_2 + 0,03X_3 + 0,17X_4 + 0,18X_5 + 0,05X_6 + 0,06X_7 - 0,52X_9 + 0,01X_{11} - 0,48X_{12} + 0,47X_{13} + 0,05X_{14} - 0,02X_{15} + 0,30X_{16} \\
 &\quad 0,36 \quad -0,37 \quad 0,05 \quad 0,27 \quad 0,28 \quad 0,08 \quad 0,10 \quad -0,81 \quad 0,01 \quad -0,76 \quad -0,74 \quad 0,08 \quad 0,03 \quad 0,48 \\
 Y_4 &= -0,17X_1 - 0,19X_2 - 0,21X_3 + 0,11X_4 + 0,32X_5 - 0,02X_6 + 0,10X_7 - 0,19X_9 - 0,77X_{11} + 0,22X_{12} + 0,15X_{13} - 0,26X_{14} + 0,04X_{15} + 0,06X_{16} \\
 &\quad -0,15 \quad -0,17 \quad -0,19 \quad 0,09 \quad 0,29 \quad -0,02 \quad 0,09 \quad -0,17 \quad -0,68 \quad 0,19 \quad 0,13 \quad -0,23 \quad 0,03 \quad 0,05
 \end{aligned}$$



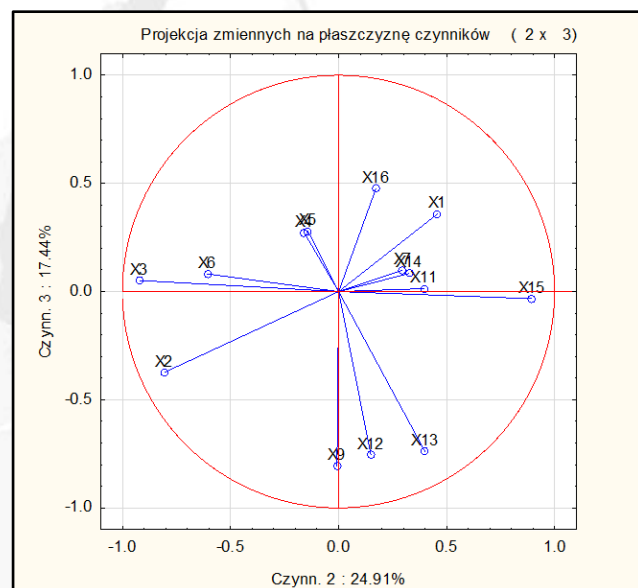
Rysunek 8. Projekcja zmiennych na płaszczyznę czynników 1x2. Opracowanie własne w programie Statistica.



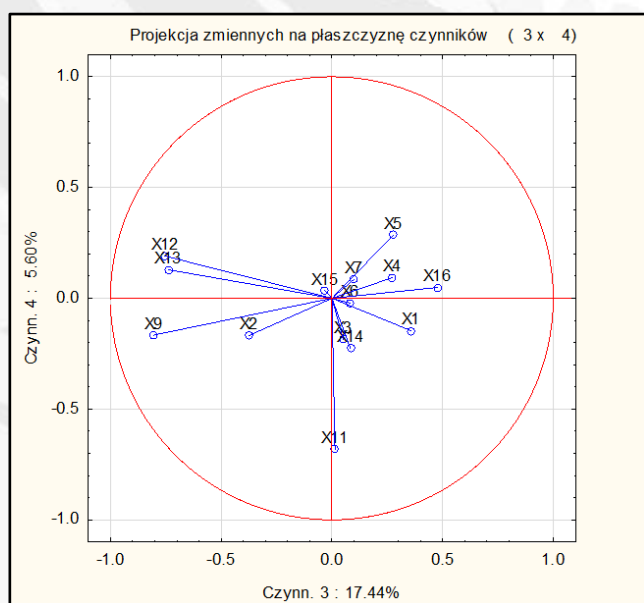
Rysunek 9. Projekcja zmiennych na płaszczyznę czynników 1x3. Opracowanie własne w programie Statistica.



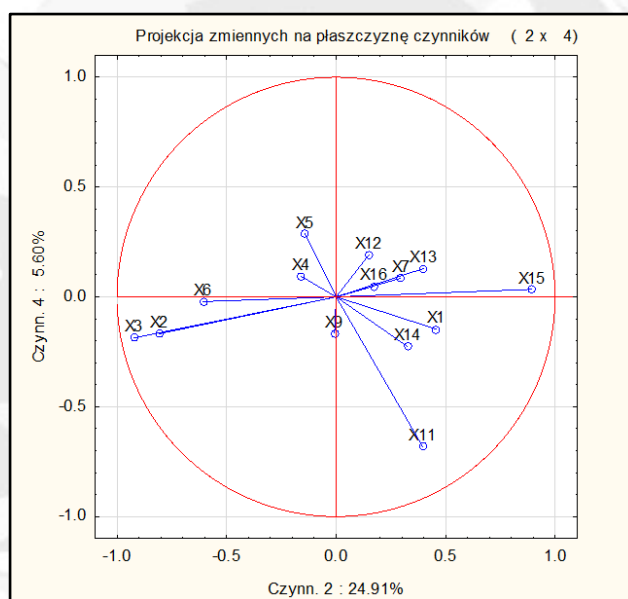
Rysunek 10. Projekcja zmiennych na płaszczyznę czynników 1x4. Opracowanie własne w programie Statistica.



Rysunek 11. Projekcja zmiennych na płaszczyznę czynników 2x3. Opracowanie własne w programie Statistica.

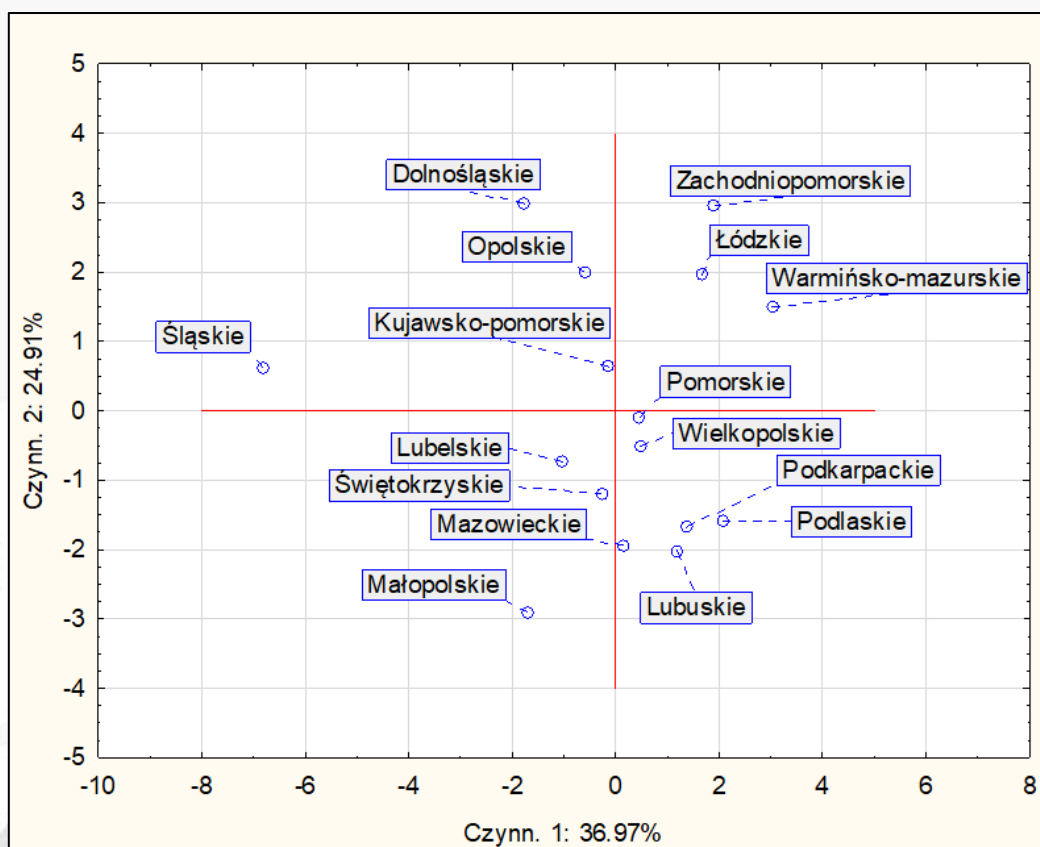


Rysunek 12. Projekcja zmiennych na płaszczyznę czynników 3x4. Opracowanie własne w programie Statistica.



Rysunek 13. Projekcja zmiennych na płaszczyznę czynników 2x4. Opracowanie własne w programie Statistica.

Na rysunkach 8 – 13 zostały przedstawione projekcje zmiennych na płaszczyznę czynników. Projekcja polega na rzutowaniu danych na płaszczyznę czynników, która jest definiowana przez najważniejsze kierunki zmienności w danych. Kierunki te są określone przez wektory własne macierzy kowariancji lub korelacji, które są obliczane wcześniej w analizie głównych składowych. Wzięto pod uwagę wszystkie możliwe kombinacje projekcji dwóch czynników na siebie spośród składowych 1 do 4. Najmocniej prezentuje się rysunek 8 czyli projekcja zmienna na płaszczyznę czynników 1x2 ponieważ przedstawia dwie pierwsze składowe, które reprezentują łącznie ponad 60% wariancji zmiennych pierwotnych. Wektory reprezentujące zmienne pierwotne zazwyczaj znajdują się na brzegach jednostkowego koła, co oznacza, że są dobrze reprezentowane przez pierwsze dwie składowe główne, które tworzą układ współrzędnych. Na wykresie analizujemy również kąt pomiędzy wektorami poszczególnych zmiennych. Mały kąt pomiędzy zmiennymi X4 i X5 świadczy o tym, że tworzą one jednorodną grupę w jednej ćwiartce układu, co jest wynikiem ich negatywnej korelacji. Zauważyliśmy, że ta grupa jest głównie reprezentowana przez pierwszą składową, ponieważ wartości współrzędnych tej grupy są wyższe dla pierwszej składowej. Wniosek z tego wynika, że wykres doskonale obrazuje brak związku między zmienną X1 (powierzchnią użytków rolnych w ha na 1000 km). Analogiczne obserwacje tylko dla innych zmiennych można wyciągnąć w przypadku analizy pozostałych projekcji. Warto jednak w tym miejscu zwrócić uwagę na projekcje zmiennych na płaszczyznę czynników np. 2x3, gdzie zmienne są porzucane po całej osi. Można się dopatrzeć zmiennych zgrupowanych, jednak w tym przypadku większość wektorów nie znajduje się przy brzegach koła co wskazuje na umiarkowanie dobrą reprezentację przez obie składowe. Wraz z kolejnymi projekcjami gdzie procentowa reprezentacja wariancji zmiennych pierwotnych jest coraz niższa można też zauważyć coraz mniej zmiennych znajdujących się na obrzeżach koła.

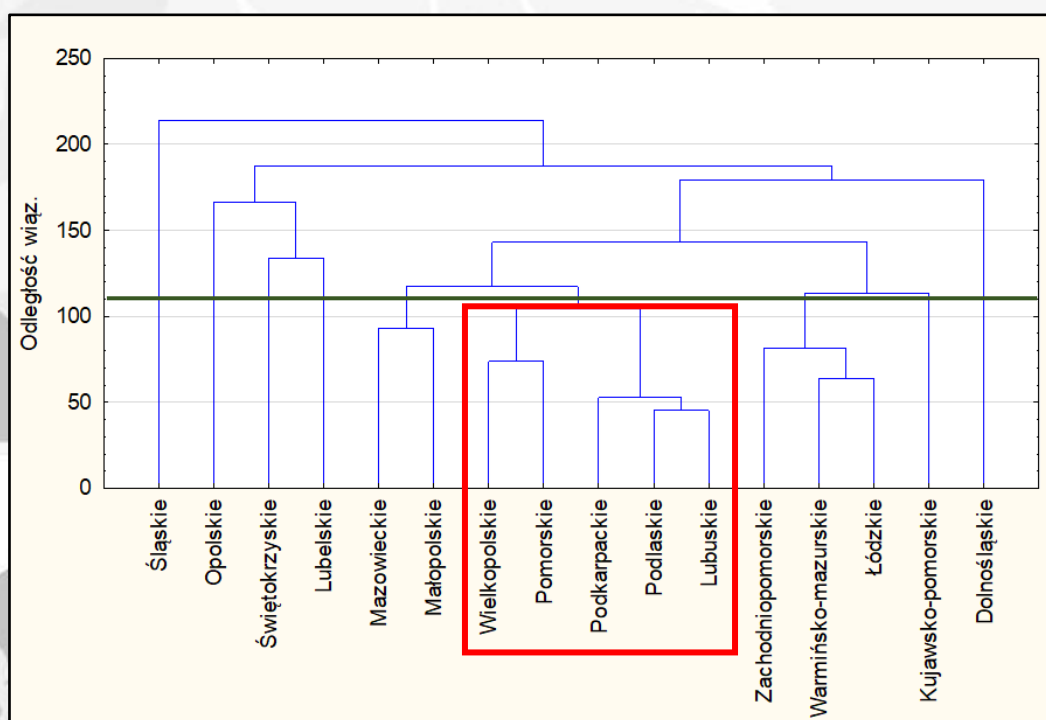


Rysunek 14. Rzut przypadków na płaszczyznę czynników 1x2. Opracowanie własne w programie Statistica.

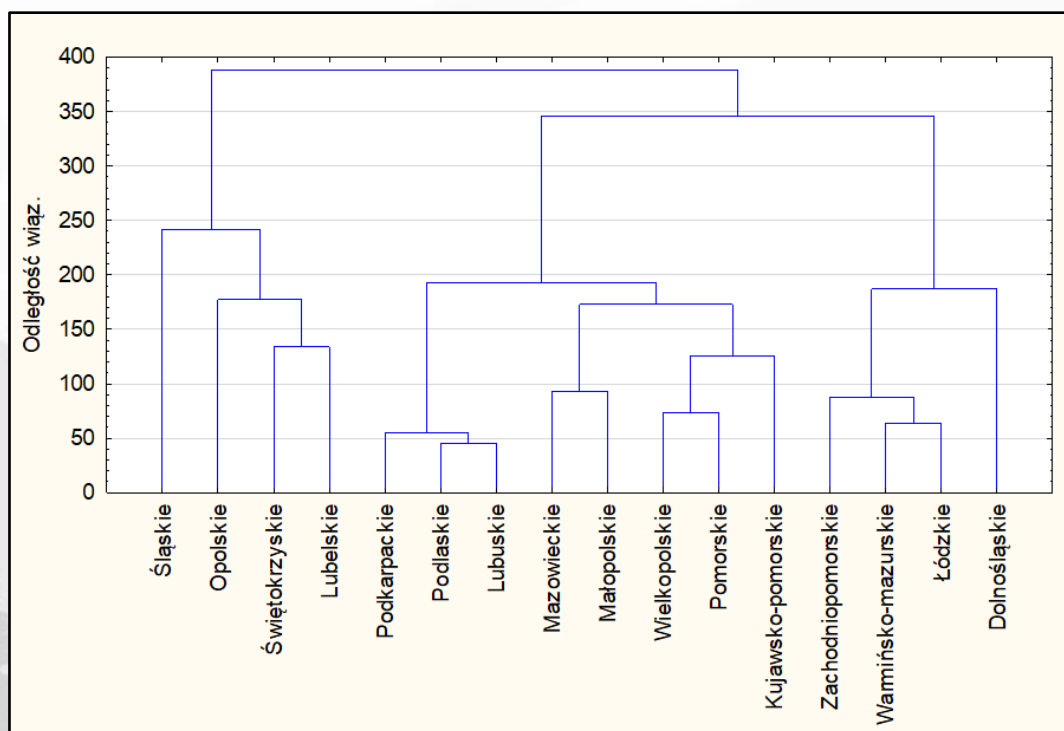
Rzut przypadków na płaszczyznę czynników ilustruje, jak poszczególne województwa wpływają na kształtowanie się składowych. Na wykresie możemy zauważyć wyjątkowy, odosobniony przypadek województwa Śląskiego, które ma największy wpływ na kształt pierwszej składowej którą można określić jako „przemysłową” czyli tej zwracającej największy wpływ zmiennych X4, X5, X7 i X14 czyli odpowiednio gęstość zaludnienia, powierzchnie budynków niemieszkalnych, długość linii kolejowych oraz powierzchnie dzikich wysypisk oraz średni wpływ na kształt drugiej. Potwierdza to również wcześniejsze obserwacje oraz wyniki analizy metodami Wzorca, TOPSIS i Sumy Rang. Z drugiej strony wykresu, można zauważyć województwa, które mają mały wpływ na pierwszą i drugą składową czyli województwa Zachodniopomorskie, Łódzkie oraz Warmińsko-mazurskie. Są to województwa które wcześniej również wypadały generalnie lepiej niż gorzej jeżeli chodzi o mały wpływ ingerencji człowieka na strukturę krajobrazową względem innych województw.

Analiza skupień

Analiza skupień, inaczej grupowanie, jest techniką uczenia nienadzorowanego, której celem jest znalezienie naturalnych grup, klastrów lub zespołów w danych. Celem tej analizy jest redukcja wielkości danych, ujawnienie ukrytych struktur oraz zrozumienie istotnych zależności pomiędzy cechami. W tej analizie zastosowano dwie metody grupowania hierarchicznego: metodę Warda i metodę średnich połączeń.



Rysunek 15. Diagram drzewa metodą Warda. Opracowanie własne w programie Statistica.



Rysunek 16. Diagram drzewa metodą średnich połączeń. Opracowanie własne w programie Statistica.

Analizując powyższe diagramy na pierwszy rzut oka ciężko wskazać konkretne grupy województw podobnych do siebie. Znając jednak wyniki poprzednich analiz, na diagramie drzewa metodą Warda można zauważyć stosunkowo dobre pogrupowanie województw: Wielkopolskiego, Pomorskiego, Podkarpackiego, Podlaskiego i Lubuskiego, które mimo wszystko charakteryzowały się najlepszymi wynikami jeżeli chodzi o naruszenie społeczne, infrastrukturalne i środowiskowe spośród wszystkich województw. We wskazanym dendrogramie został ustawiony punkt odcięcia na odległość około 120, co doprowadziło do utworzenia 8 klastrow. Na podstawie analizy dendrogramu stworzono mapę klastrow widoczną na rysunku 17.



Rysunek 17. Mapa klastrow na podstawie wyników obliczeń wykonanych przy wykorzystaniu Metody Wzorca z wagami. Opracowanie własne w programie Excel.

Wnioski

Celem niniejszej analizy było zidentyfikowanie województwa, którego krajobraz naturalny został najmniej naruszony działalnością człowieka. Do badania zostały w większości wybrane zmienne, które jednoznacznie wskazywały na nieobojętny, a często bardzo negatywny wpływ dla naturalności środowiska. Analiza pozwoliła wyłonić jednocześnie województwa najbardziej dzikie, ale też takie, których środowisko było zdecydowanie najbardziej naruszone (co niekoniecznie znaczy że w najgorszym stanie).

Województwa, które wypadły najlepiej to Podkarpackie, Podlaskie, Wielkopolskie, Lubuskie i Pomorskie. Gdyby jednak należało wybrać tylko jedno województwo, to wskazanie Podkarpackiego wydaje się najbardziej stosowne. Jest to jednostka administracyjna, która w każdym rodzaju metod rankingowych była w czołówce i najczęściej plasowała się na pierwszym miejscu, jednak ciężko było się doszukiwać bardzo znaczących różnic względem pozostałych województw, które również „dobrze” wypadły. Na rysunku 13, gdzie przedstawiany był rzut przypadków na płaszczyznę czynników 1 i 2 widać, że bardzo bliskie sąsiedztwo Lubuskiego i Podlaskiego w badaniu, gdzie dwie pierwsze składowe, reprezentowały razem ponad 60% wariancji zmiennych pierwotnych sugeruje, że do pierwszego miejsca należy podchodzić z dystansem i raczej warto w tym kontekście mówić bardziej o całym klastrze ustalonym podczas przeprowadzania Analizy Skupień niż wskazywać jednostkę. Najistotniejsze w kontekście tego badania okazały się zmienne mówiące o gęstości zaludnienia, powierzchni i liczbie budynków, długości linii kolejowych, powierzchniach dzikich wysypisk oraz powierzchni i wycince lasów.

Jest możliwym natomiast wskazanie województwa które wypadło najgorzej. Zdecydowanie najbardziej naruszoną działalnością ludzką jest województwo Śląskie, które wypadło najgorzej w każdym jednym badaniu i które wykazywało największy wpływ na kształtowanie pierwszej składowej. Wynik taki nie jest dużym zaskoczeniem, z racji tego, że jak powszechnie wiadomo jest to województwo, które charakteryzuje się największą aglomeracją miejską oraz licznymi ośrodkami przemysłowymi, wyjątkowo jak na skalę całej polski. Bardzo słabo wypadły też województwa Dolnośląskie, Świętokrzyskie oraz Zachodniopomorskie.

Ogólnie wyniki przeprowadzonej analizy nie są obiektywnie satysfakcjonujące. Różnice w wynikach jakie można zaobserwować w najlepszej piątce nie są bardzo duże, a i środek stawki w generalnym rozrachunku wypada bardzo podobnie względem siebie. Co prawda udało się wskazać województwo, którego środowisko naturalne zostało naruszone w największym stopniu, jednak pierwotne założenia były zupełnie odwrotne. Można z dużą dozą pewności stwierdzić, że jeżeli ktoś szuka azylu w możliwie najbardziej dziewiczym krajobrazowo otoczeniu to niech porzuci pomysł by zaczynać poszukiwania od województwa śląskiego, jednak żeby wskazać jednoznacznie województwo od którego powinien zacząć poszukiwania, to z tym może być ciężiej. Niesatysfakcjonujące wyniki najprawdopodobniej wynikają z nieidealnie dobranych do badania zmiennych, lub po prostu z bardzo podobnego poziomu

urbanizacji większości polskich województw. Niemniej jednak podobną analizę można byłoby przeprowadzić w większej skali dla krajów Europy lub świata. Różnice na poziomie międzynarodowym mogłyby być bardziej uwypuklone.

Największą ciekawostką jaką udało się zaobserwować podczas zbierania danych był pobór wody na potrzeby gospodarki narodowej i ludności w ciągu roku na mieszkańca w województwie Świętokrzyskim. Wartość wskazuje na prawie dwukrotnie większy pobór względem drugiego w tej kategorii województwa Zachodniopomorskiego i prawie 10 razy większy niż chociażby województwo Pomorskie. Nie udało się niestety znaleźć potwierdzonych informacji co wpływa na tak duże wykorzystanie zasobów wodnych w tym regionie.



Rysunek 18. Kadr z filmu Wyspa Psów, reż. Wes Anderson [9]

Bibliografia

[1] Alok R. U. i Sunil J. V. (2015), *Urbanization and Mental Health: A Review of the Literature and a Public Health Response*, International Journal of Environmental Research and Public Health, 12(7), 7497-7519.

[2] Corey, G. (2013), *The Psychological Benefits of Urban Nature: Exploring the Value of Community Gardens*, Journal of Environmental Psychology, 36, 1-10.

[3] Neale, C., & Stroud, D. (2019), *The Desire for Nature: Examining the Pull of the Natural World*, Journal of Environmental Psychology, 60, 1-11.

[4] Braund, T. (2019), *For the Love of Nature: Exploring the Importance of Species Diversity and Micro-Variables Associated with Favorite Outdoor Places*, Frontiers in Psychology, 10, 1-11.

[5] Chen, Y. (2021), *Remote Work in the COVID-19 Era: Exploring the Work-Life Balance and Job Satisfaction of Remote Workers*, International Journal of Human Resource Management, 32(15), 2765-2782.

[6] Tolstoj, L. (1899), *Zmartwychwstanie*, Książka i Wiedza (1957), Warszawa, Część I, Rozdział I.

[7] Baza danych Głównego Urzędu Statystycznego, dostęp 01.02.2023 pod adresem: <https://bdl.stat.gov.pl/bdl/dane/podgrup/temat>

[8] Khairulina, A. I. (2008), *Metody porządkowania w badaniach regionalnych i regionalistycznych*, Wydawnictwo Naukowe PWN, Warszawa, Polska.

[9] Anderson, W. (2018), *Wyspa Psów*, Imperial – Cinepix, USA.

Wszystkie ilustracje krajobrazów wykorzystane w pracy pochodzą ze źródeł stockowych serwisu Vectostock, dostęp 05.02.2023 w bazie dostępnej pod adresem: <https://www.vectorstock.com/>

