Obraz zawierający tekst, Czcionka, logo, symbol

Opis wygenerowany automatycznie

Zastosowanie metod wielowymiarowej analizy porównawczej do oceny stanu środowiska w województwie dolnośląskim

Mikołaj Ziółkowski

Wydział Zarządzania

Informatyka i Ekonometria, rok 3

Kraków, 2024r.

1. Dane:

Badanie przeprowadzono na podstawie danych statystycznych o stanie i ochronie środowiska w 30 powiatach województwa dolnośląskiego w 2022 r., udostępnionych na stronie Głównego Urzędu Statystycznego(GUS) . Uwzględniono 9 zmiennych.

|  |  |  |
| --- | --- | --- |
| Zmienne | OPIS | Charakter zmiennej |
| X1 | Pobór wody na potrzeby gospodarki narodowej i ludności[dam3/os] | destymulanta |
| X2 | Zużycie wody na potrzeby gospodarki narodowej i ludności[dam3/os] | destymulanta |
| X3 | Ścieki przemysłowe i komunalne/ludność korzystająca z oczyszczalni ścieków | destymulanta |
| X4 | Oczyszczalnie ścieków[szt/os] | stymulanta |
| X5 | emisja i redukcja zanieczyszczeń powietrza | destymulanta |
| X6 | ochrona przyrody i krajobrazu — powierzchnia[ha] | stymulanta |
| X7 | ochrona przyrody i krajobrazu — pomniki przyrody[szt/ha] | stymulanta |
| X8 | Odpady wytworzone [t/os] | destymulanta |
| X9 | odpady dotychczas składowane i tereny ich składowania[t/ha] | destymulanta |

Zmienne powstały na podstawie Banku Danych Lokalnych(BDL) w następujący sposób:

1. Zmienna X1 -> Pobór wody na potrzeby gospodarki narodowej i ludności[dam^3] podzielono przez liczbę ludności danego powiatu.
2. Zmienna X2 -> Zużycie wody na potrzeby gospodarki narodowej i ludności[dam^3] podzielono przez liczbę ludności danego powiatu.
3. Zmienna X3 -> Ścieki przemysłowe i komunalne podzielono przez ludność korzystającą z oczyszczalni ścieków.
4. Zmienna X4 -> Ścieki przemysłowe i komunalne[szt] podzielono przez liczbę ludności danego powiatu.
5. Zmienna X5 -> [(Emisja zanieczyszczeń pyłowych) - (redukcja zanieczyszczeń pyłowych)]+[( Emisja zanieczyszczeń gazowych) – (redukcja zanieczyszczeń gazowe)].
6. Zmienna X6 -> ochrona przyrody i krajobrazu — powierzchnia[ha] podzielona przez powierzchnię[ha] danego powiatu.
7. Zmienna X7 -> ochrona przyrody i krajobrazu — pomniki przyrody[szt] podzielona przez powierzchnię[ha] danego powiatu.
8. Zmienna X8 -> odpady wytworzone[t] podzielone prze liczbę ludności danego powiatu.
9. Zmienna X9 -> odpady dotychczas składowane[t] podzielone przez tereny ich składowania[ha].

Zmienne przeskalowano w celu możliwości lepszego porównywania wartości pomiędzy powiatami, które różnią się pod wieloma czynnikami, m.in.: powierzchnią, ludnością, etc.

* 1. Dane w formie tabeli

Tabela o wymiarach 11x31.

Kolumny: L.P ,Poiwaty i zmienne X1,X2,X3,X4,X5,X6,X7,X8,X9

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| L.P | Powiaty | X1 | X2 | X3 | X4 | X | X6 | X7 | X8 | X9 |
| 1 | bolesławiecki | 0,0920 | 0,0779 | 0,0460 | 0,0001 | 36813,0 | 0,0920 | 0,0014 | 4,44 | 99896,1 |
| 2 | dzierżoniowski | 0,0327 | 0,0352 | 0,0349 | 0,0001 | 33846,0 | 0,1788 | 0,0025 | 0,19 | 0,0 |
| 3 | głogowski | 0,1905 | 0,0423 | 0,0367 | 0,0002 | 385885,0 | 0,0279 | 0,0008 | 13,13 | 123618,4 |
| 4 | górowski | 0,0675 | 0,0492 | 0,0576 | 0,0002 | 3477,0 | 0,7059 | 0,0011 | 0,20 | 0,0 |
| 5 | jaworski | 0,0527 | 0,0445 | 0,0356 | 0,0001 | 20288,0 | 0,2629 | 0,0020 | 2,26 | 0,0 |
| 6 | karkonoski | 0,0972 | 0,0544 | 0,0465 | 0,0002 | 55834,0 | 0,2625 | 0,0009 | 0,23 | 0,0 |
| 7 | kamiennogórski | 0,2500 | 0,0337 | 0,0417 | 0,0003 | 13592,0 | 0,1603 | 0,0007 | 1,80 | 57823,9 |
| 8 | kłodzki | 0,0614 | 0,0379 | 0,0527 | 0,0001 | 33815,0 | 0,3755 | 0,0011 | 1,83 | 355713,0 |
| 9 | legnicki | 0,1733 | 0,1701 | 0,0381 | 0,0002 | 37829,0 | 0,1201 | 0,0018 | 1,30 | 0,0 |
| 10 | lubański | 0,0546 | 0,0395 | 0,0521 | 0,0002 | 21489,0 | 0,0225 | 0,0032 | 0,78 | 0,0 |
| 11 | lubiński | 0,0579 | 0,4076 | 0,0394 | 0,0002 | 148655,0 | 0,0212 | 0,0011 | 74,03 | 315998,6 |
| 12 | lwówecki | 0,0324 | 0,0262 | 0,0306 | 0,0002 | 845,0 | 0,1326 | 0,0007 | 0,28 | 0,0 |
| 13 | milicki | 2,1749 | 2,1592 | 0,0350 | 0,0001 | 0,0 | 0,6776 | 0,0007 | 0,00 | 0,0 |
| 14 | oleśnicki | 0,0826 | 0,0720 | 0,0421 | 0,0001 | 32540,0 | 0,1046 | 0,0006 | 0,08 | 0,0 |
| 15 | oławski | 0,0715 | 0,0637 | 0,0478 | 0,0001 | 55686,0 | 0,0015 | 0,0006 | 2,21 | 0,0 |
| 16 | polkowicki | 0,0471 | 0,0388 | 0,0366 | 0,0002 | 266730,0 | 0,2669 | 0,0003 | 346,70 | 340259,2 |
| 17 | strzeliński | 0,0887 | 0,0684 | 0,0360 | 0,0002 | 72323,0 | 0,1465 | 0,0004 | 2,05 | 0,0 |
| 18 | średzki | 0,0594 | 0,0467 | 0,0400 | 0,0002 | 53537,0 | 0,0074 | 0,0004 | 0,59 | 0,0 |
| 19 | świdnicki | 0,0622 | 0,0459 | 0,0461 | 0,0001 | 114831,0 | 0,0743 | 0,0025 | 0,41 | 0,0 |
| 20 | trzebnicki | 0,3192 | 0,3059 | 0,0426 | 0,0002 | 386,0 | 0,2401 | 0,0020 | 0,22 | 0,0 |
| 21 | wałbrzyski | 0,0976 | 0,0291 | 0,0311 | 0,0002 | 837,0 | 0,3557 | 0,0031 | 0,53 | 0,0 |
| 22 | wołowski | 0,2464 | 0,2314 | 0,1893 | 0,0001 | 104730,0 | 0,1179 | 0,0004 | 1,71 | 45587,7 |
| 23 | wrocławski | 0,0891 | 0,1109 | 0,0723 | 0,0001 | 533717,0 | 0,1097 | 0,0013 | 2,59 | 82036,4 |
| 24 | ząbkowicki | 0,0471 | 0,0365 | 0,0403 | 0,0003 | 0,0 | 0,1585 | 0,0021 | 4,18 | 226882,8 |
| 25 | zgorzelecki | 0,3383 | 0,3089 | 0,0386 | 0,0002 | 9220266,0 | 0,0014 | 0,0005 | 23,07 | 0,0 |
| 26 | złotoryjski | 0,0515 | 0,0360 | 0,0351 | 0,0003 | 166,0 | 0,0705 | 0,0012 | 11,43 | 0,0 |
| 27 | \*miasto Jelenia Góra | 0,0645 | 0,0466 | 0,0877 | 0,0000 | 88607,0 | 0,1991 | 0,0016 | 0,12 | 0,0 |
| 28 | \*miasto Legnica | 0,1019 | 0,0510 | 0,0487 | 0,0000 | 154914,0 | 0,0027 | 0,0096 | 4,52 | 107505,4 |
| 29 | \*miasto Wrocław | 0,1262 | 0,1076 | 0,0573 | 0,0000 | 1132852,0 | 0,0240 | 0,0039 | 0,87 | 229923,1 |
| 30 | \*miasto Wałbrzych | 0,0052 | 0,0459 | 0,0353 | 0,0000 | 268684,0 | 0,1694 | 0,0076 | 1,59 | 166027,3 |

\*miasto na prawach powiatu

1. Statystyki opisowe:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 |
| Min | 0.005 | 0.026 | 0.03 | 1.187e-05 | 0 | 0.0014 | 0.0003 | 0 | 0 |
| 1st Q | 0.0554 | 0.039 | 0.036 | 8.523e-05 | 15266 | 0.039 | 0.0007 | 0.3137 | 0 |
| Mediana | 0.077 | 0.048 | 0.041 | 1.517e-04 | 45683 | 0.126 | 0.00117 | 1.65 | 0 |
| Średnia | 0.175 | 0.161 | 0.049 | 1.410e-04 | 429772 | 0.169 | 0.0019 | 16.779 | 717 |
| 3rd Q | 0.12 | 0.1 | 0.049 | 1.951e-04 | 140199 | 0.229 | 0.002 | 3.782 | 105603 |
| Max | 2.174 | 2.159 | 0.189 | 2.697e-04 | 9220266 | 0.7059 | 0.0096 | 346.7036 | 355713 |
| Odch.Stan | 0.39 | 0.39 | 0.03 | 0 | 1675711 | 0.18 | 0 | 63.83 | 112766.62 |
| Kurtoza | 20.82 | 20.11 | 15 | -0.92 | 22.25 | 2.57 | 5.58 | 20.81 | 0.51 |

Zmienne X1,X2 i X3 mają stosunkowo małe wartości mediany i średniej, ale duży rozstęp, co sugeruje, że dane są bardzo zróżnicowane, bądź istnieją znaczące outliery. Kurtoza w tych zmiennych jest wysoka, co sugeruje na rozkład z wąskim szczytem, a tym samym występowanie dużej ilości ekstremalnych wartości. X5 Ma dużą różnicę między wartością minimalną a maksymalną, co wskazuje na znaczną zmienność danych. Średni jest zdecydowanie większa niż mediana, co oznacz, że prawdopodobnie istnieją duże wartości odstające. X9 i X10 mają duże różnice między wartościami max i min, co wskazuje na dużą zmienność danych i obecność potencjalnych wartości odstających.

1. Korelacja:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *X1* | *X2* | *X3* | *X4* | *X5* | *X6* | *X7* | *X8* | *X9* |
| *X1* | 1,000 |  |  |  |  |  |  |  |  |
| *X2* | 0,977 | 1,000 |  |  |  |  |  |  |  |
| *X3* | -0,032 | -0,032 | 1,000 |  |  |  |  |  |  |
| *X4* | 0,054 | 0,032 | -0,226 | 1,000 |  |  |  |  |  |
| *X5* | 0,067 | 0,060 | -0,052 | 0,150 | 1,000 |  |  |  |  |
| *X6* | 0,509 | 0,483 | -0,046 | 0,055 | -0,217 | 1,000 |  |  |  |
| *X7* | -0,147 | -0,142 | -0,101 | -0,535 | -0,091 | -0,159 | 1,000 |  |  |
| *X8* | -0,074 | -0,037 | -0,102 | 0,278 | 0,035 | 0,044 | -0,164 | 1,000 |  |
| *X9* | -0,160 | -0,086 | -0,027 | 0,034 | -0,064 | -0,048 | 0,128 | 0,528 | 1,000 |

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

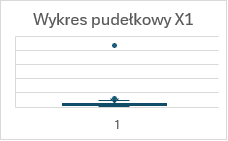
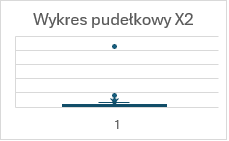
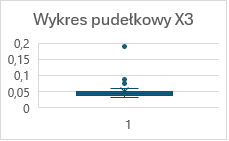
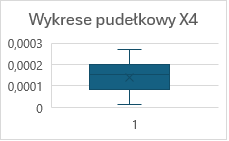
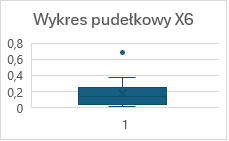
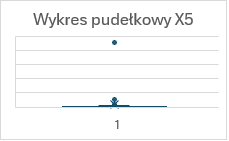
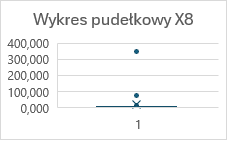
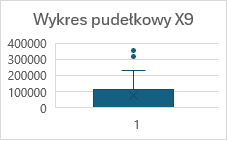
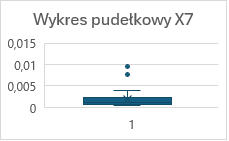
Na podstawie wyników korelacji można zauważyć, iż istnieje wysoka zależność(0,98) między X1 i X2. Ta wysoka korelacja wskazuje, iż w dalszym badaniu należy rozważyć usunięcie jednej ze zmiennych X1/X2, o ile inne, późniejsze modyfikacje nie zmienią odpowiednio wartości korelacji.

1. Współczynnik zmienności:

|  |  |
| --- | --- |
| Zmienne | Wsp. zmiennosci |
| X1 | 221.72 |
| X2 | 242.264 |
| X3 | 59.299 |
| X4 | 52.739 |
| X5 | 389.91 |
| X6 | 103.305 |
| X7 | 110.35 |
| X8 | 380.44 |
| X9 | 157.256 |

Bardzo wysokie współczynniki zmienności oznaczają, że wartości zmiennych są bardzo rozporoszone wokół średniej, co może wynikać z obecności outlierów. W zmiennych X5 i X8 współczynnik wskazuje na ekstremalne wartości, co prawdopodobnie jest wynikiem związanym z nielicznymi bardzo dużymi wartościami odstającymi.

1. Wartości odstające(outliery)
   1. Boxploty

  
  
Na podstawie wykresów pudełkowych można wywnioskować, iż w danych istnieją znaczące wartości odstające w większości przypadków. Jedynie zmienna X4 nie posiada wartości odstających.

* 1. Test Grubbsa

Używam testu Grubbsa do wykrywania wartości odstających w próbach

- hipoteza zerowa H0: W próbie nie ma wartości odstających

-hipoteza alternatywna H1: W próbie występują wartości odstające

Jeśli p-value < 0.05, to odrzucamy H0

|  |  |
| --- | --- |
| Zmienne | p-value |
| X1 | 0 |
| X2 | 0 |
| X3 | 5.8280047454673e-11 |
| X4 | 1 |
| X5 | 0 |
| X6 | 0.01241 |
| X7 | 0.000189 |
| X8 | 0 |
| X9 | 0.1183 |

Dla zmiennych X1,X2,X3,X5,X6,X7,X8 odrzucamy H0 i przyjmujemy, że występują wartości odstające. Dla Zmiennej X4, p-value-1, co oznacza, że nie ma wartości odstających. Dla Zmiennej X9 nie ma wystarczających dowodów na występowanie wartości odstających.

1. Podsumowanie wstępnej analizy.

Przez wzgląd na istnienie znaczących wartości odstających wśród większości ze zmiennych, decyduję się na zmianę wartości outlierów na wartość końca ,,wąsa” tzn. na najwyższy wynik, który nie jest outlierem w odrębie danej zmiennej. Wartość końca wąsa obliczana jest ze wzroru: , gdzie IQR = Q3 – Q1 , Q3-kwantyl trzeci, Q1-kwantyl pierwszy.  
  
6.1 Zamiana outlierów

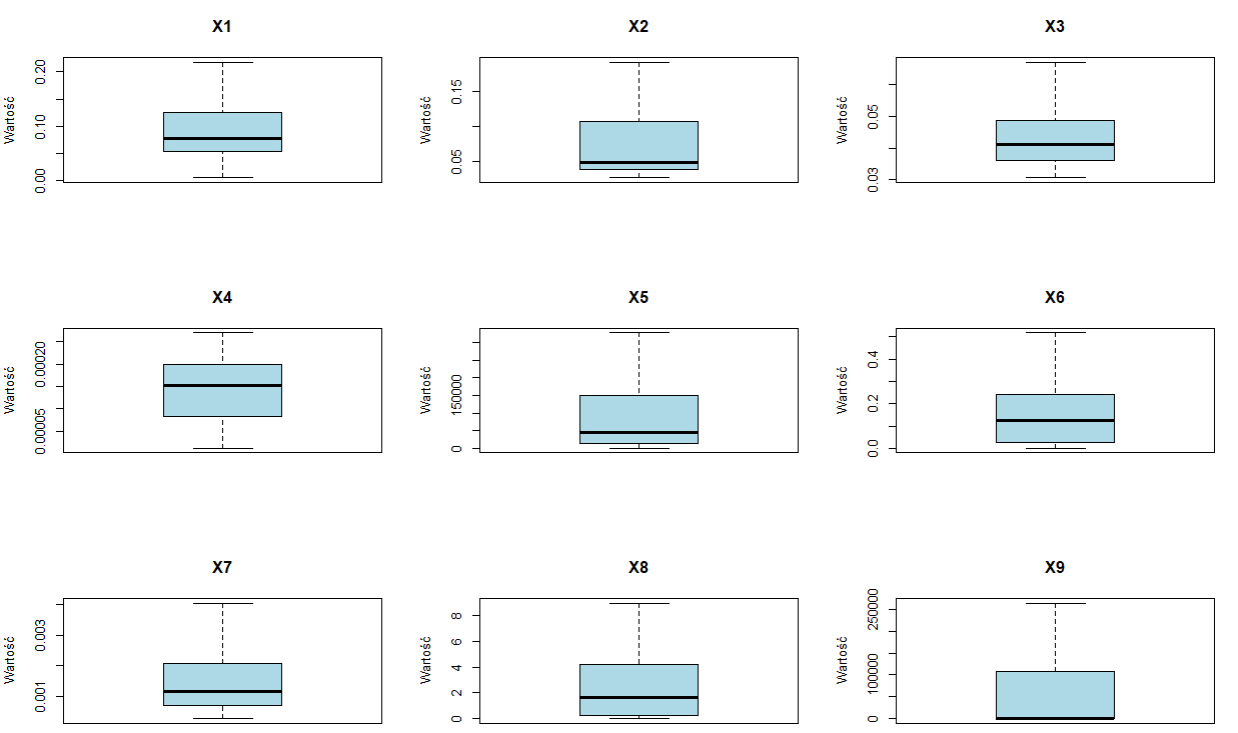
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| L.P | Powiaty | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** |
| 1 | bolesławiecki | 0,092 | 0,078 | 0,046 | 0,000080 | 36813,000 | 0,092 | 0,001 | 4,443 | 99896,057 |
| 2 | dzierżoniowski | 0,033 | 0,035 | 0,035 | 0,000083 | 33846,000 | 0,179 | 0,002 | 0,193 | 0,000 |
| 3 | głogowski | 0,190 | 0,042 | 0,037 | 0,000198 | 327598,500 | 0,028 | 0,001 | 8,986 | 123618,399 |
| 4 | górowski | 0,067 | 0,049 | 0,058 | 0,000153 | 3477,000 | 0,517 | 0,001 | 0,195 | 0,000 |
| 5 | jaworski | 0,053 | 0,045 | 0,036 | 0,000105 | 20288,000 | 0,263 | 0,002 | 2,257 | 0,000 |
| 6 | karkonoski | 0,097 | 0,054 | 0,047 | 0,000213 | 55834,000 | 0,263 | 0,001 | 0,228 | 0,000 |
| 7 | kamiennogórski | 0,217 | 0,034 | 0,042 | 0,000270 | 13592,000 | 0,160 | 0,001 | 1,802 | 57823,864 |
| 8 | kłodzki | 0,061 | 0,038 | 0,053 | 0,000114 | 33815,000 | 0,375 | 0,001 | 1,835 | 264007,616 |
| 9 | legnicki | 0,173 | 0,170 | 0,038 | 0,000164 | 37829,000 | 0,120 | 0,002 | 1,301 | 0,000 |
| 10 | lubański | 0,055 | 0,039 | 0,052 | 0,000154 | 21489,000 | 0,023 | 0,003 | 0,780 | 0,000 |
| 11 | lubiński | 0,058 | 0,192 | 0,039 | 0,000203 | 148655,000 | 0,021 | 0,001 | 8,986 | 264007,616 |
| 12 | lwówecki | 0,032 | 0,026 | 0,031 | 0,000185 | 845,000 | 0,133 | 0,001 | 0,280 | 0,000 |
| 13 | milicki | 0,217 | 0,192 | 0,035 | 0,000140 | 0,000 | 0,517 | 0,001 | 0,000 | 0,000 |

Pierwsze 13 wierszy po zmianie:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *X1* | *X2* | *X3* | *X4* | *X5* | *X6* | *X7* | *X8* | *X9* |
| X1 | 1,00 |  |  |  |  |  |  |  |  |
| X2 | 0,66 | 1,00 |  |  |  |  |  |  |  |
| X3 | 0,09 | 0,14 | 1,00 |  |  |  |  |  |  |
| X4 | 0,23 | 0,04 | -0,44 | 1,00 |  |  |  |  |  |
| X5 | 0,12 | 0,18 | 0,20 | -0,23 | 1,00 |  |  |  |  |
| X6 | 0,03 | -0,04 | -0,04 | 0,06 | -0,39 | 1,00 |  |  |  |
| X7 | -0,25 | -0,19 | 0,05 | -0,55 | 0,13 | -0,11 | 1,00 |  |  |
| X8 | 0,06 | 0,11 | -0,24 | 0,45 | 0,48 | -0,36 | -0,23 | 1,00 |  |
| X9 | -0,19 | -0,02 | 0,07 | 0,00 | 0,45 | -0,06 | 0,16 | 0,45 | 1,00 |

6.2 korelacja po zmianie:  
  
Po modyfikacji outlierów, korelacja również się zmieniła, żadna wartość w macierzy nie wskazuje na wartość powyżej 0.9. Nie ma już potrzeby rozważania usunięcia jednej z zmiennych X1 i X2, jak to było przed zmianą.

6.3 Test Grubbsa i wykresy pudełkowe



|  |  |
| --- | --- |
| Zmienne | p-value |
| X1 | 0.989 |
| X2 | 0.747 |
| X3 | 0359 |
| X4 | 1 |
| X5 | 0.605 |
| X6 | 0.112 |
| X7 | 0.386 |
| X8 | 0.571 |
| X9 | 0.468 |

Brak wartości odstających, ich wpływ został zniwelowany.

1. Porządkowanie liniowe

W kolejnym etapie badania przechodzę do porządkowania liniowego. Do analizy wykorzystuje 3 metody: Hellwiga, Topsis oraz Standaryzowanych Sum.

! ustalenie wag zmiennych — przyjęto takie same wagi dla wszystkich zmiennych

7.1 Wyniki poszczególnych metod porządkowania liniowego:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Hellwiga | | | Topsis | | | StandaryzowanychSum | | |
| Ranking | | | Ranking | | | Ranking | | |
| L.P | Powiaty | WSK | L.P | Powiaty | WSK | L.P | Powiaty | WSK |
| 8 | kłodzki | 0,523332 | 21 | wałbrzyski | 0,777031 | 8 | kłodzki | 1 |
| 21 | wałbrzyski | 0,449676 | 4 | górowski | 0,726133 | 21 | wałbrzyski | 0,970858 |
| 24 | ząbkowicki | 0,433668 | 5 | jaworski | 0,682775 | 2 | dzierżoniowski | 0,943255 |
| 2 | dzierżoniowski | 0,423075 | 2 | dzierżoniowski | 0,681823 | 12 | lwówecki | 0,91904 |
| 14 | oleśnicki | 0,407974 | 6 | karkonoski | 0,661427 | 4 | górowski | 0,856297 |
| 6 | karkonoski | 0,406938 | 12 | lwówecki | 0,648615 | 19 | świdnicki | 0,844954 |
| 4 | górowski | 0,399603 | 10 | lubański | 0,642647 | 5 | jaworski | 0,840941 |
| 5 | jaworski | 0,397333 | 27 | miasto Jelenia Góra | 0,601675 | 24 | ząbkowicki | 0,830072 |
| 19 | świdnicki | 0,394597 | 19 | świdnicki | 0,601608 | 14 | oleśnicki | 0,803324 |
| 12 | lwówecki | 0,368037 | 13 | milicki | 0,597006 | 6 | karkonoski | 0,772136 |
| 1 | bolesławiecki | 0,353842 | 14 | oleśnicki | 0,591334 | 10 | lubański | 0,763803 |
| 20 | trzebnicki | 0,335549 | 17 | strzeliński | 0,580658 | 20 | trzebnicki | 0,762922 |
| 10 | lubański | 0,329044 | 20 | trzebnicki | 0,576189 | 18 | średzki | 0,685015 |
| 26 | złotoryjski | 0,287795 | 7 | kamiennogórski | 0,571972 | 26 | złotoryjski | 0,683453 |
| 17 | strzeliński | 0,28537 | 18 | średzki | 0,570081 | 1 | bolesławiecki | 0,628951 |
| 9 | legnicki | 0,285133 | 24 | ząbkowicki | 0,566764 | 17 | strzeliński | 0,611183 |
| 18 | średzki | 0,281345 | 26 | złotoryjski | 0,558099 | 13 | milicki | 0,603635 |
| 13 | milicki | 0,235768 | 9 | legnicki | 0,555028 | 9 | legnicki | 0,585023 |
| 15 | oławski | 0,229755 | 8 | kłodzki | 0,552676 | 15 | oławski | 0,522427 |
| 7 | kamiennogórski | 0,220355 | 30 | miasto Wałbrzych od 2013 | 0,539167 | 7 | kamiennogórski | 0,51611 |
| 23 | wrocławski | 0,215969 | 15 | oławski | 0,534913 | 30 | miasto Wałbrzych od 2013 | 0,430221 |
| 16 | polkowicki | 0,19415 | 1 | bolesławiecki | 0,502451 | 23 | wrocławski | 0,363085 |
| 30 | miasto Wałbrzych od 2013 | 0,160956 | 28 | miasto Legnica | 0,492465 | 16 | polkowicki | 0,343279 |
| 28 | miasto Legnica | 0,145648 | 22 | wołowski | 0,443347 | 27 | miasto Jelenia Góra | 0,342425 |
| 11 | lubiński | 0,13949 | 23 | wrocławski | 0,417026 | 11 | lubiński | 0,277527 |
| 29 | miasto Wrocław | 0,102596 | 29 | miasto Wrocław | 0,407397 | 28 | miasto Legnica | 0,255059 |
| 3 | głogowski | 0,099814 | 16 | polkowicki | 0,406581 | 3 | głogowski | 0,206215 |
| 27 | miasto Jelenia Góra | 0,068254 | 25 | zgorzelecki | 0,353803 | 29 | miasto Wrocław | 0,123508 |
| 25 | zgorzelecki | 0,029886 | 3 | głogowski | 0,343871 | 25 | zgorzelecki | 0,101216 |
| 22 | wołowski | 0,003284 | 11 | lubiński | 0,331416 | 22 | wołowski | 0 |

7.2 Ranking(od 1 do 30) według poszczególnych metod:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Hellwig | Topsis | StandaryzowanychSum |
| kłodzki | 1 | 19 | 1 |
| wałbrzyski | 2 | 1 | 2 |
| ząbkowicki | 3 | 16 | 8 |
| dzierżoniowski | 4 | 4 | 3 |
| oleśnicki | 5 | 11 | 9 |
| karkonoski | 6 | 5 | 10 |
| górowski | 7 | 2 | 5 |
| jaworski | 8 | 3 | 7 |
| świdnicki | 9 | 9 | 6 |
| lwówecki | 10 | 6 | 4 |
| bolesławiecki | 11 | 22 | 15 |
| trzebnicki | 12 | 13 | 12 |
| lubański | 13 | 7 | 11 |
| złotoryjski | 14 | 17 | 14 |
| strzeliński | 15 | 12 | 16 |
| legnicki | 16 | 18 | 18 |
| średzki | 17 | 15 | 13 |
| milicki | 18 | 10 | 17 |
| oławski | 19 | 21 | 19 |
| kamiennogórski | 20 | 14 | 20 |
| wrocławski | 21 | 25 | 22 |
| polkowicki | 22 | 27 | 23 |
| miasto Wałbrzych | 23 | 20 | 21 |
| miasto Legnica | 24 | 23 | 26 |
| lubiński | 25 | 30 | 25 |
| miasto Wrocław | 26 | 26 | 28 |
| głogowski | 27 | 29 | 27 |
| miasto Jelenia Góra | 28 | 8 | 24 |
| zgorzelecki | 29 | 28 | 29 |
| wołowski | 30 | 24 | 30 |

Uzyskane wyniki charakteryzują się istotnym zróżnicowaniem, przy czym zaskakujący jest fakt, iż większa zbieżność dotyczy rankingów pomiędzy wzorcowymi metodami a bez wzorcowymi. W szczególności dotyczy to metod Hellwiga i Stand.Sum., co również wynika z poniższej korelacji. Można zauważyć, że w metodzie Hellwiga oraz Standaryzowanych Sum w pierwszej czwórce rankingu znajdują się powiaty: kłodzki(1), wałbrzyski(2), ząbkowicki i dzierżoniowski.

7.3 Korelacja Tau-Kendalla pomiędzy rankingami metod

|  |  |  |  |
| --- | --- | --- | --- |
|  | Hellwig | Topsis | StandSum |
| Hellwig | 1 |  |  |
| Topsis | 0.526 | 1 |  |
| Stand.Sum | 0.844 | 0.609 | 1 |

7.4 Grupowanie według średniej  
  
a) obliczenia :

|  |  |  |  |
| --- | --- | --- | --- |
|  | Hellwiga | Topsis | StandaryzowanychSum |
| średnia | 0,273607784 | 0,558088348 | 0,586197817 |
| odchyenie | 0,136803892 | 0,106133756 | 0,283571825 |
| suma | 0,410411676 | 0,664222104 | 0,869769642 |
| różnica | 0,136803892 | 0,451954592 | 0,302625992 |
| GR.I | r>=0,4104 | r>=0,6642 | r>=0,8697 |
| GR.II | 0,2736<=r<0,4104 | 0,558<=r<0,6642 | 0,5861<=r<0,8697 |
| GR.III | 0,1368<=r<0,2736 | 0,4519<=r<0,558 | 0,3026<=r<0,5861 |
| GR.IV | r<0,1368 | r<0,4519 | r<0,3026 |

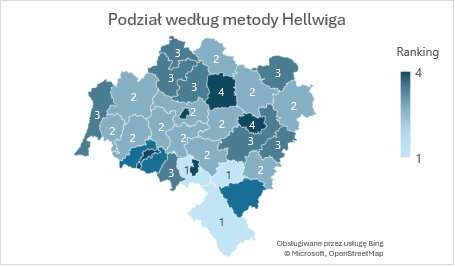
b) Klasyfikacja powiatów województw

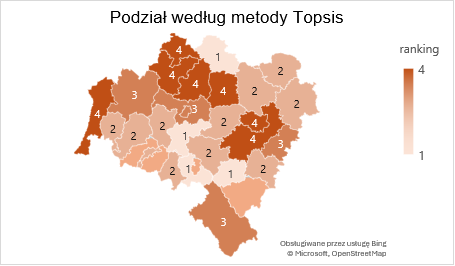
W klasie pierwszej znalazły się powiaty wyróżniające się najlepszym stanem ochrony środowiska ([klasa I]zgodna klasyfikacja w przypadku powiatów: wałbrzyski, dzierżoniowski), a w czwartej – najgorszym ([klasa IV]zgodna klasyfikacja w przypadku powiatów miasto Wrocław, głogowski, zgorzelecki, wołowski)

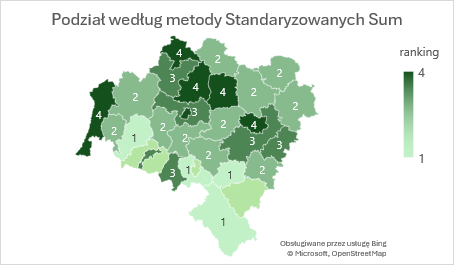
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Klasa | Hellwiga | | | Topsis | | | Standaryzowanych Sum | | |
| I | kłodzki |  | 4 | wałbrzyski | | 4 | kłodzki |  | 4 |
| wałbrzyski | |  | górowski |  |  | wałbrzyski | |  |
| ząbkowicki | |  | jaworski |  |  | dzierżoniowski | |  |
| dzierżoniowski | |  | dzierżoniowski | |  | lwówecki |  |  |
| II | oleśnicki |  | 13 | karkonoski | | 13 | górowski |  | 13 |
| karkonoski | |  | lwówecki |  |  | świdnicki |  |  |
| górowski |  |  | lubański |  |  | jaworski |  |  |
| jaworski |  |  | miasto Jelenia Góra | |  | ząbkowicki | |  |
| świdnicki |  |  | świdnicki |  |  | oleśnicki |  |  |
| lwówecki |  |  | milicki |  |  | karkonoski | |  |
| bolesławiecki | |  | oleśnicki |  |  | lubański |  |  |
| trzebnicki |  |  | strzeliński | |  | trzebnicki |  |  |
| lubański |  |  | trzebnicki |  |  | średzki |  |  |
| złotoryjski | |  | kamiennogórski | |  | złotoryjski | |  |
| strzeliński | |  | średzki |  |  | bolesławiecki | |  |
| legnicki |  |  | ząbkowicki | |  | strzeliński | |  |
| średzki |  |  | złotoryjski | |  | milicki |  |  |
| III | milicki |  | 8 | legnicki |  | 6 | legnicki |  | 7 |
| oławski |  |  | kłodzki |  |  | oławski |  |  |
| kamiennogórski | |  | miasto Wałbrzych | | | kamiennogórski | |  |
| wrocławski | |  | oławski |  |  | miasto Wałbrzych | | |
| polkowicki | |  | bolesławiecki | |  | wrocławski | |  |
| miasto Wałbrzych | | | miasto Legnica | |  | polkowicki | |  |
| miasto Legnica | |  |  |  |  | miasto Jelenia Góra | |  |
| lubiński |  |  |  |  |  |  |  |  |
| IV | miasto Wrocław | | 5 | wołowski |  | 7 | lubiński |  | 6 |
| głogowski |  |  | wrocławski | |  | miasto Legnica | |  |
| miasto Jelenia Góra | |  | miasto Wrocław | |  | głogowski |  |  |
| zgorzelecki | |  | polkowicki | |  | miasto Wrocław | |  |
| wołowski |  |  | zgorzelecki | |  | zgorzelecki | |  |
|  |  |  | głogowski |  |  | wołowski |  |  |
|  |  |  | lubiński |  |  |  |  |  |

7.5 Wizualizacja grupowania

Na mapach poniże zilustrowano klasyfikację powiatów ze względu na stan ochrony środowiska (mniej intensywny kolor oznacza klasę powiatów charakteryzujących się lepszym stanem ochrony środowiska)







Z przedstawionych powyżej ilustracji można zauważyć, iż powiaty o gorszym stanie ochrony środowiska znajdują się w północno-zachodniej części województwa dolnośląskiego, co wynika z wszystkich zastosowanych metod. Natomiast te o lepszym stanie znajdują się na południu oraz w centrum województwa.

d) wykres radarowy

Oś radialna: reprezentuje różne powiaty. Każdy z powiatów jest reprezentowany przez swoje wartości w trzech metodach.

Oś wartości: Okręgi oznaczające wartości wskaźników dla poszczególnych metod.

Zauważalny jest fakt, iż metoda Hellwiga zazwyczaj ma mniejsze wartości w porównaniu do pozostałych metod. Dla Topsis wartości są zazwyczaj wyższe niż w przypadku Hellwiga. W niektórych powiatach (np. kłodzki, wołowski, wrocławski) widzimy znacząco wyższe wartości w porównaniu do Hellwiga. Standaryzowanych sum: Podobnie jak TOPSIS, ta metoda wykazuje większe rozróżnienie między powiatami, co oznacza bardziej widoczne różnice w wynikach. Powiat Kłodzki uzyskuje wysokie wskaźniki we wszystkich trzech metodach, co oznacza, że według każdej z nich jego sytuacja jest oceniana pozytywnie. Ząbkowicki i wałbrzyski: Te powiaty również są wysoko oceniane, szczególnie w metodach TOPSIS i Standaryzowanych Sum. Powiaty takie jak polkowicki, milicki, średzki: Mają niższe wskaźniki w każdej z metod, co sugeruje, że są oceniane słabiej według wszystkich trzech algorytmów.

1. Analiza Skupień

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznieW kolejnej części badania przeprowadzam analizę skupień dwoma metodami, metodą hierarchiczną Warda oraz metodą podziałową k-średnich. Do metody k-średnich poszukuje optymalnej liczby klastrów za pomocą metody ,,Łokciowej” oraz metody „Sylwetkowej”.  
  
8.1Metoda\_Łokciowa:

Metoda ta wskazuje, iż optymalną liczbą klastrów jest liczba 4.

Obraz zawierający tekst, diagram, linia, Wykres

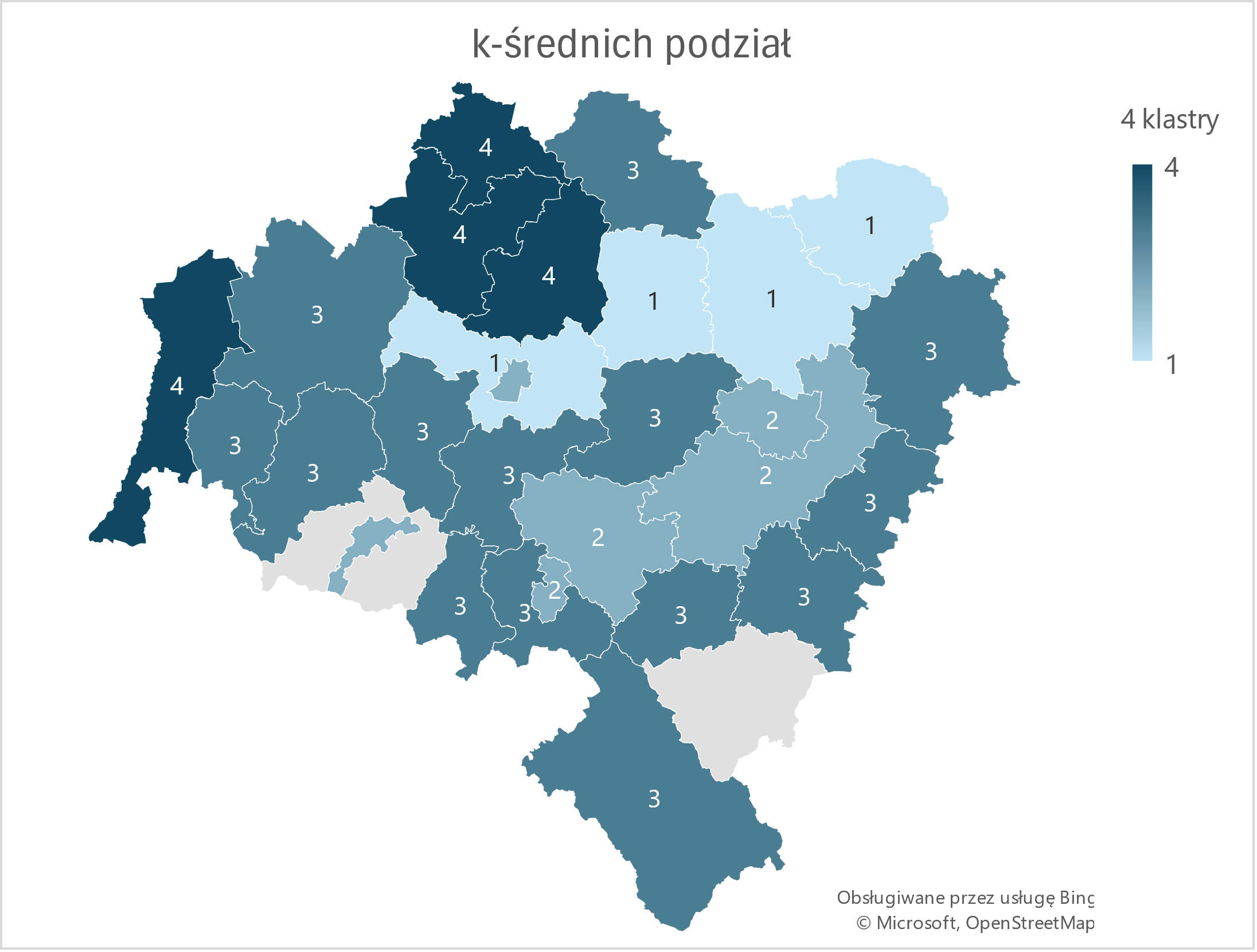
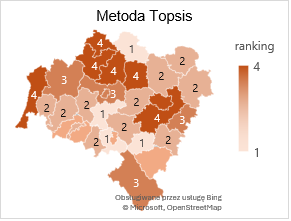
Opis wygenerowany automatycznie8.2 Metoda Sylwetkowa:

Natomiast ta metoda wskazuje, iż optymalną liczbą klastrów będzie 5.

8.3 Metoda k-średnich

Porównanie podziałów przez wzgląd na 4 i 5 klastrów, oraz wybór najbardziej optymalnego rozwiązania.

|  |  |  |
| --- | --- | --- |
| Powiat | 5 klastrów | 4 klastry |
| bolesławiecki | 5 | 3 |
| dzierżoniowski | 5 | 3 |
| głogowski | 2 | 4 |
| górowski | 1 | 3 |
| jaworski | 5 | 3 |
| karkonoski | 5 | 3 |
| kamiennogórski | 5 | 3 |
| kłodzki | 1 | 3 |
| legnicki | 4 | 1 |
| lubański | 5 | 3 |
| lubiński | 2 | 4 |
| lwówecki | 5 | 3 |
| milicki | 4 | 1 |
| oleśnicki | 5 | 3 |
| oławski | 5 | 3 |
| polkowicki | 2 | 4 |
| strzeliński | 5 | 3 |
| średzki | 5 | 3 |
| świdnicki | 5 | 2 |
| trzebnicki | 4 | 1 |
| wałbrzyski | 5 | 3 |
| wołowski | 4 | 1 |
| wrocławski | 3 | 2 |
| ząbkowicki | 5 | 3 |
| zgorzelecki | 2 | 4 |
| złotoryjski | 5 | 3 |
| miasto Jelenia Góra | 1 | 2 |
| miasto Legnica | 3 | 2 |
| miasto Wrocław | 3 | 2 |
| miasto Wałbrzych | 3 | 2 |

Porównując otrzymane wyniki podziału z rankingami z porządkowania liniowego dochodzę do wniosku, iż podział na 4 klastry jest lepszym podziałem ponieważ widać pewne zależności pomiędzy podziałem na 4 klastry a rankingiem wg. Metod: Hellwiga, Topsis, StandaryzowanychSum.  
  
8.4 Podział na 4 klastry:

W powyższym podziale metodą k-średnich, można za pomocą ilustracji zauważy pewne podobieństwa do metody Topsis w podziale powiatów, co pozytywnie wpływa na przebieg badania. Jednak wciąż istnieje dość spora różnica między metodami porządkowania liniowego a metodą k-średnich, przykładowo powiat Kłodzki dla metod Hellwiga i Standaryzowanych Sum jest na szczycie rankingu, natomiast w k-średnich w klastrze numer 3.

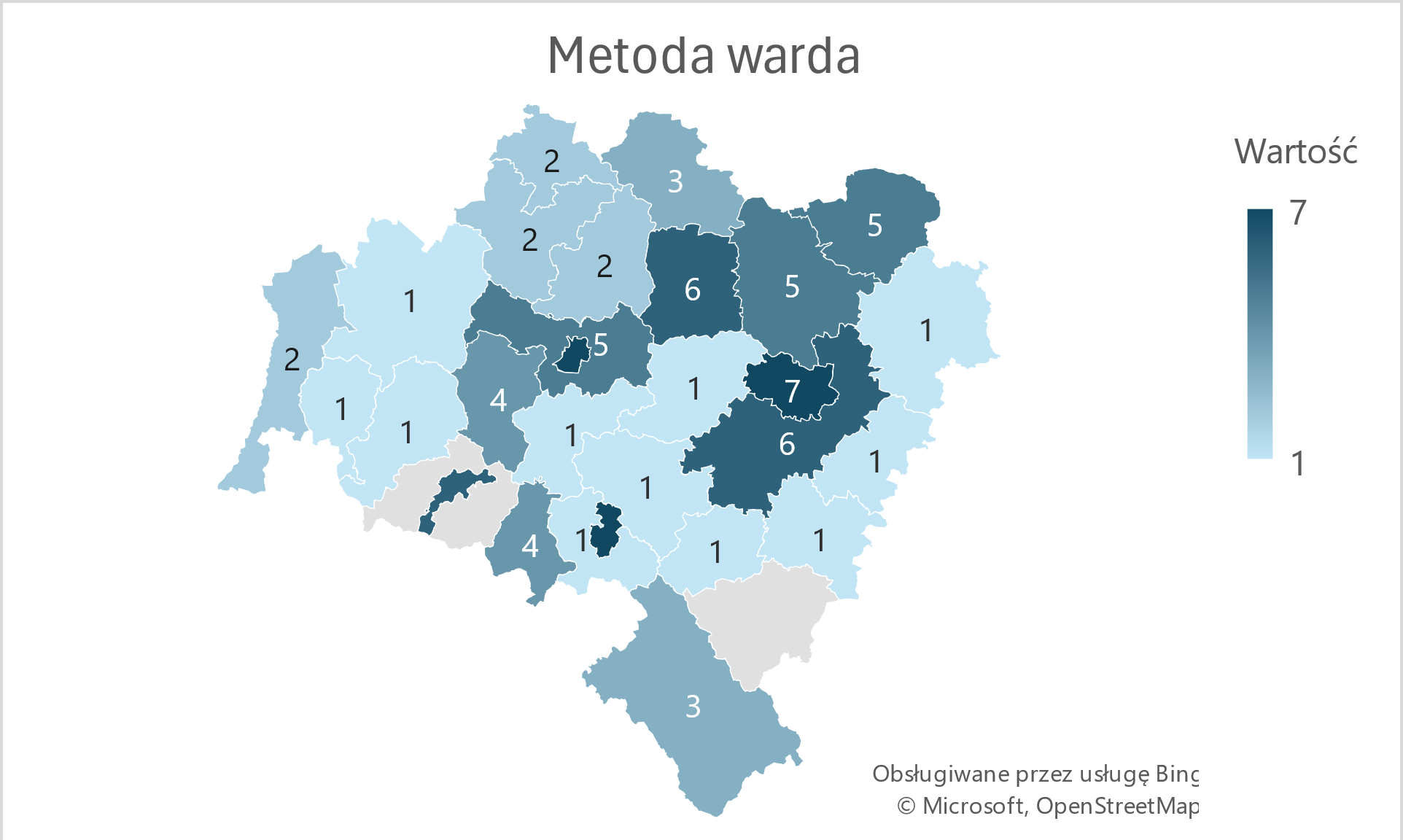
8.5 Hierarchiczna metoda Warda

Podział metodą Warda dzieli powiaty na 7 obszarów, co obrazuje poniższy dendrogram

Obraz zawierający tekst, diagram, linia, Równolegle

Opis wygenerowany automatycznie

|  |  |
| --- | --- |
| Powiat | Podział |
| bolesławiecki | 1 |
| dzierżoniowski | 1 |
| głogowski | 2 |
| górowski | 3 |
| jaworski | 1 |
| karkonoski | 3 |
| kamiennogórski | 4 |
| kłodzki | 3 |
| legnicki | 5 |
| lubański | 1 |
| lubiński | 2 |
| lwówecki | 1 |
| milicki | 5 |
| oleśnicki | 1 |
| oławski | 1 |
| polkowicki | 2 |
| strzeliński | 1 |
| średzki | 1 |
| świdnicki | 1 |
| trzebnicki | 5 |
| wałbrzyski | 1 |
| wołowski | 6 |
| wrocławski | 6 |
| ząbkowicki | 4 |
| zgorzelecki | 2 |
| złotoryjski | 4 |
| Jelenia Góra | 6 |
| miasto Legnica | 7 |
| miasto Wrocław | 7 |
| miasto Wałbrzych | 7 |



Na podstawie powyższego podziału oraz mapy można zauważy podobieństwo do poprzednich metod. Mianowicie powiaty ,,lepsze,, w centrum i na południu województwa, natomiast nieco gorsze na północy. Zgodność z innymi metodami np.: z metodami porządkowania liniowego występuje przykładowo w powiecie wołowskim oraz w powiecie miast Wrocław (jako gorsze) oraz zgodność z nieco lepszymi jak świdnicki, kłodzkim, jaworski.

8.6 Statystyki w grupach i ich analiza

średnia

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| klaster | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 |
| 1 | 0.066 | 0.049 | 0.0402 | 0.0000123 | 40276 | 0.125 | 0.00158 | 1.26 | 9801 |
| 2 | 0.128 | 0.116 | 0.0378 | 0.000213 | 2676460 | 0.0794 | 0.000680 | 8.99 | 162908 |
| 3 | 0.0754 | 0.0472 | 0.0523 | 0.000160 | 31042 | 0.385 | 0.00104 | 0.753 | 88003 |
| 4 | 0.105 | 0.0354 | 0.0390 | 0.000267 | 4586 | 0.130 | 0.00134 | 4.99 | 94902 |
| 5 | 0.202 | 0.185 | 0.0386 | 0.000154 | 12738 | 0.292 | 0.0149 | 0.506 | 0 |
| 6 | 0.124 | 0.117 | 0.0669 | 0.0000737 | 173645 | 0.142 | 0.00112 | 1.47 | 42541 |
| 7 | 0.0778 | 0.0682 | 0.0471 | 0.0000212 | 250399 | 0.0654 | 0.00399 | 2.33 | 167819 |

Odchylenie standardowe

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 |
| 1 | 0.0226 | 0.0178 | 0.00716 | 0.0000416 | 33128 | 0.109 | 0.00112 | 1.34 | 30120 |
| 2 | 0.0881 | 0.0874 | 0.00140 | 0.000143 | 84357 | 0.126 | 0.000361 | 0 | 127181 |
| 3 | 0.0192 | 0.00842 | 0.00556 | 0.0000498 | 26288 | 0.127 | 0.000114 | 0.937 | 152425 |
| 4 | 0.0969 | 0.00148 | 0.00352 | 0.00000424 | 7800 | 0.0513 | 0.000669 | 3.66 | 117898 |
| 5 | 0.0253 | 0.0127 | 0.00383 | 0.0000130 | 21730 | 0.203 | 0.000690 | 0.697 | 0 |
| 6 | 0.0819 | 0.0729 | 0 | 0.0000591 | 133571 | 0.0494 | 0.000646 | 1.25 | 41103 |
| 7 | 0.0640 | 0.0343 | 0.0111 | 0.0000103 | 87782 | 0.0907 | 0.0000816 | 1.93 | 61229 |

Wnioski:

X1:destymulanta  
Klastry 1,3,7 mają niższe wartości średniego poboru wody (X1), co jest pozytywne, bo X1 jest dystymulantą. Powiaty w tych klastrach średnio zużywają mniej wody na mieszkańca. Klastry 5 i 2 mają wyższe wartości– oznacza to, że powiaty w tych klastrach mają wyższe zużycie wody, co jest negatywne z punktu widzenia tej zmiennej.

X2:destymulanta

Klastry 1,3,4 mają niższe wartości zużycia wody, co wpływa pozytywnie .Natomiast klastry 2,5 i 6 mają najwyższe średnie zużycie wody, co jest bardziej negatywne.

X3:destymulanta

Klastry 2,4,5 mają niską wartość, co oznacz lepsze wyniki w zakresie ścieków na osobę. Klastry 3 i 6 mają najwyższą wartość, co sugeruje gorszy wpływ na ranking.

X4:stymulanta

Klaster 4 ma najwyższą wartość, co oznacza, że w powiatach tych jest lepsza infrastruktura oczyszczalni ścieków na osobę. Klaster 7 ma najniższą wartość, co jest negatywne.

X5:destymulanta

Klaster 4 ma stosunkowo niską wartość w porównaniu do pozostałych, co oznacza że te powiaty mają niski poziom emisji zanieczyszczeń. Natomiast problem z emisją zanieczyszczeń przypuszczalnie znajduje się w klastrach 7 i 2, ponieważ tam są najwyższe wartości.

X6:stymulanta

Klastry 2 i 3 mają największą wartość, co wskazuje na obecność większych obszarów objętych ochroną przyrody i krajobrazu w tych powiatach niż w pozostałych. Klaster 7 ma najniższą wartość, co może sugerować, że w tych powiatach jest mniej obszarów chronionych.

X7:stymulanta

Klaster 5 ma najwyższą wartość, a tym samym w tych powiatach znajduje się więcej pomników przyrody niż w pozostałych. Klaster 2 ma bardzo niską wartość, a tym samym te powiaty mają najmniej pomników przyrody.

X8:destymulanta

Klastry 1,3 i 5 mają niskie wartości, co jest pozytywne i oznacza mniejszą produkcje odpadów na mieszkańca. Wartość w klastrach 2 i 4 sugeruje, iż powiaty te mają większy problem z odpadami od pozostałych

X9:destymulanta

Klaster 1 ma niską wartość, a tym samym oznacza to mniejsze składowiska odpadów na jednostkę terenu. Większe problemy związane z odpadami obejmują klastry 2 i 7, gdzie wartości są dość duże w porównaniu do reszty klastrów.

Podsumowanie Klastrów:

Klaster 1 : Można zauważyć, iż powiaty w tym klastrze mają stosunkowo niskie wartości dla destymulant(w szczególności dla X1,X2,X8,X9, co oznacza lepsze wyniki. Odchylenie standardowe jest stosunkowo niskie ,co sugeruje małą zmienność między powiatami w tym klastrze.

Klaster 7: „Najgorszy” Powiaty w tym klastrze charakteryzują się stosunkowo wysokimi średnimi wartościami dla destymulant, co oznacza większe problemy np.: wysoka emisja zanieczyszczeń, dużą produkcją odpadów. Jednocześnie mają niższe wartości dla niektórych stymulant np.: oczyszczalnie ścieków, co wskazuje na niedostateczną infrastrukturę.

Klaster 4 : Wyróżnił się najlepiej pod względem oczyszczalni ścieków na osobę X4. Jednocześnie powiaty te mają także najniższą emisję zanieczyszczeń.

Klaster 5 i 2: klastry te charakteryzują się wysokim poborem(X1) i zużyciem wody(X2), co jest niekorzystne. Klaster 2 dodatkowo ma stosunkowo wysoką wartość zanieczyszczeń.  
  
Niskie odchylenia standardowe w klastrach (szczególnie w klastrach 1, 3 i 4) sugerują, że powiaty w tych klastrach są bardziej jednorodne. Wysokie odchylenia standardowe, szczególnie w zmiennych takich jak X9 (odpady składowane) i X5 (emisja zanieczyszczeń), wskazują na dużą różnorodność w zarządzaniu odpadami i emisją zanieczyszczeń wśród powiatów. Przykładowo, klaster 2 i 7 są wyraźnie bardziej zróżnicowane, co może wskazywać na istnienie zarówno powiatów z poważnymi problemami środowiskowymi, jak i tych, które radzą sobie lepiej.