

Analiza zmienności preferencji wyborczych w Polsce

Norbert Gwiazda

Rozpoczęcie: 04.03.2025r. Zakończenie: 6 czerwca 2025

1 Wprowadzenie

W Polsce wybory wygrywa ten, kto uzyska najwięcej głosów, jednak nie wszystkie regiony kraju charakteryzują się stałością preferencji wyborczych. W niektórych powiatach poparcie dla poszczególnych partii potrafi znacząco się zmieniać pomiędzy kolejnymi wyborami. W niniejszym badaniu postanowiłem przeanalizować różne miary zmienności poparcia oraz ich korelacje z wybranymi cechami powiatów, aby sprawdzić, czy można wyróżnić obszary o szczególnie dużej lub małej zmienności preferencji wyborczych.

2 Definicja miary zmienności

W analizie zmienności preferencji wyborczych na poziomie powiatów posłużymy nam miara zmienności.

Niech:

- P oznacza zbiór partii politycznych,
- T będzie zbiorem wszystkich wyborów objętych analizą,
- $\nu_{p,t}$ oznacza poparcie partii $p \in P$ w wyborach $t \in T$.

W pierwszej kolejności wyprowadzimy zmienność poparcia jednej partii w powiecie:

$$\Delta_p = \sum_{t=1}^{n-1} |\nu_{p,t+1} - \nu_{p,t}| \quad (1)$$

Zatem, zmienność wszystkich partii w powiecie będzie sumą:

$$\Delta_{calkowita} = \sum_{p \in P} \Delta_p \quad (2)$$

Aby zachować czytelność miary, będzie ona znormalizowana do przedziału $\langle 0, 1 \rangle$:

$$Z_{znormalizowana} = \frac{\Delta_{calkowita} - \min(\Delta_{calkowita})}{\max(\Delta_{calkowita}) - \min(\Delta_{calkowita})} \quad (3)$$

Aby urozmaicić analizę, wyprowadzimy następujące miary zmienności:

- Miara zmienności w wyborach samorządowych: T będzie zbiorem wyborów samorządowych,
- Miara zmienności w wyborach innych niż samorządowe: T będzie zbiorem wyborów innych niż samorządowe,

- Miara zmienności dwóch głównych partii w wyborach samorządowych: T będzie zbiorem wyborów samorządowych, a P będzie zbiorem { PiS, KO }
- Miara zmienności dwóch głównych partii w wyborach innych niż samorządowe: T będzie zbiorem wyborów innych niż samorządowe, a P będzie zbiorem { PiS, KO }
- Miara zmienności mniejszych partii w wyborach samorządowych: T będzie zbiorem wyborów samorządowych, a P będzie zbiorem { Lewica, TD, Konfederacja }

Dodatkowo wprowadzimy również miarę zmienności różnicy poparcia dwóch głównych partii w wyborach, która będzie wyglądać następująco:

$$\Delta_{PiS-KO} = \sum_{t=1}^{n-1} |\nu_{PiS,t+1} - \nu_{KO,t+1}| - |\nu_{PiS,t} - \nu_{KO,t}| \quad (4)$$

$$\Delta_{calkowita} = \sum_{p \in P} \Delta_p \quad (5)$$

$$Z_{znormalizowana} = \frac{\Delta_{calkowita} - \min(\Delta_{calkowita})}{\max(\Delta_{calkowita}) - \min(\Delta_{calkowita})} \quad (6)$$

Również i ją podzielimy na:

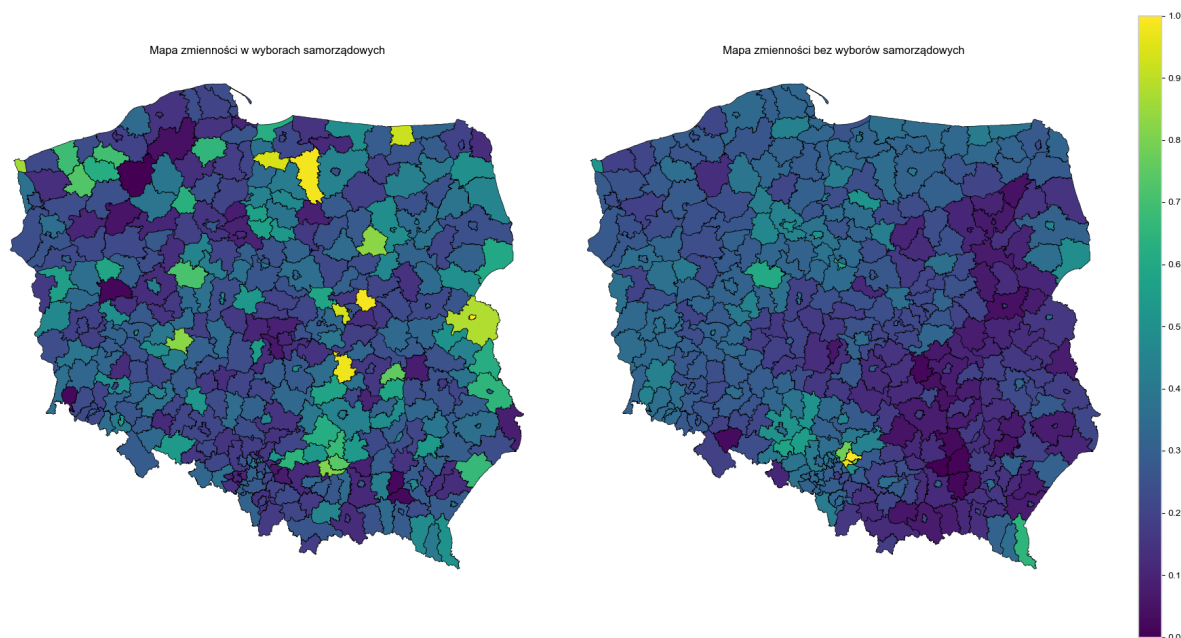
- Miara zmienności różnicy poparcia dwóch głównych partii w wyborach samorządowych: T będzie zbiorem wyborów samorządowych.
- Miara zmienności różnicy poparcia dwóch głównych partii w wyborach innych niż samorządowe: T będzie zbiorem wyborów innych niż samorządowe.

Co więcej, aby móc uśrednić miary zmienności w wyborach samorządowych i niesamorządowych, uśredniamy ich wartości w miarach „skumulowanych”.

$$Z_{skumulowana} = \frac{Z_{samorzadowa} + Z_{niesamorzadowa}}{2} \quad (7)$$

3 Analiza zmienności na mapach

Analizę zmienności na mapach zaczniemy od zmienności w wyborach samorządowych (Rys 1.)

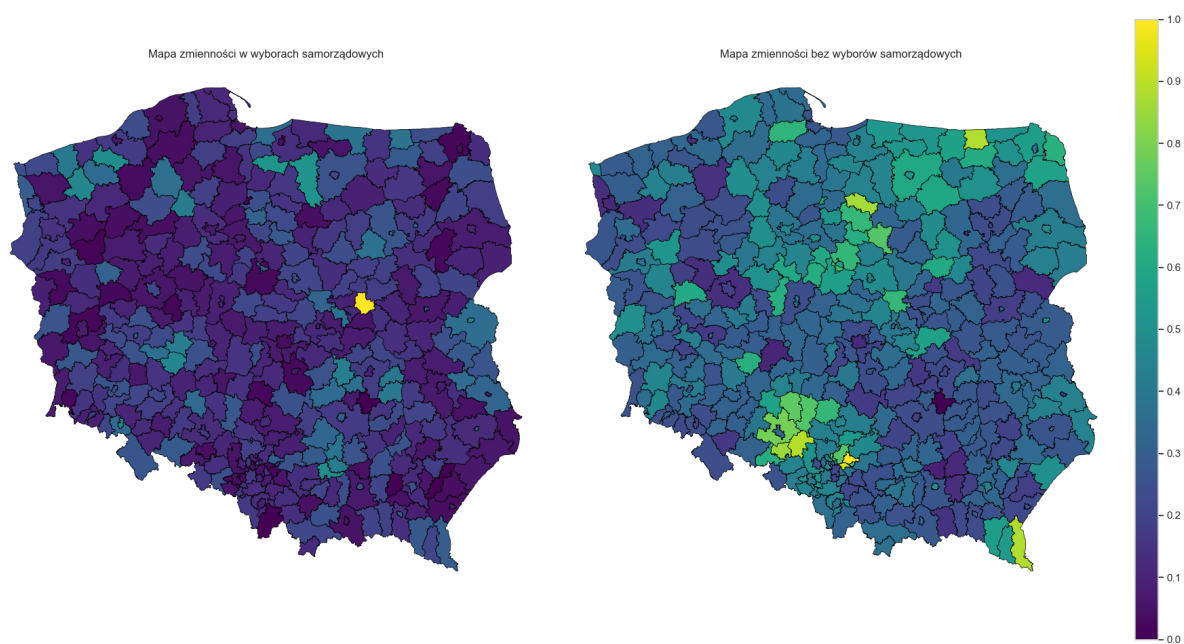


Rysunek 1: Mapy zmienności

Dostrzegalne jest to, że w wyborach samorządowych poparcie jest dużo bardziej zmienne niż w reszcie typów wyborów. Może to wynikać z lokalnego charakteru wyborów samorządowych, w których lokalny komitet lub znany samorządowiec potrafi przełamać typowe podziały partyjne, zbierając szerokie poparcie w regionie.

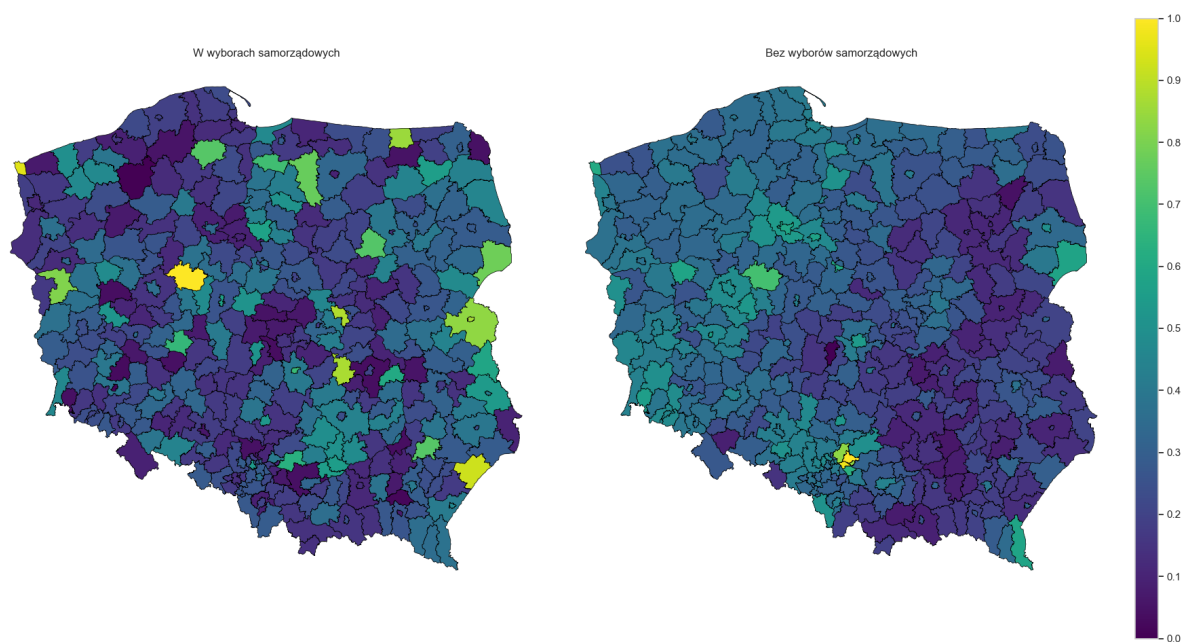
W wyborach samorządowych wysoka zmienność poparcia występuje w powiatach: sztumskim, ostródzkim, żyrardowskim, przysuskim, Warszawie i Białej Podlaskiej. Niska zmienność poparcia obserwowana jest natomiast w powiatach: pabianickim, szczecineckim, bytowskim, hrubieszowskim, lubańskim, nowotomyskim, strzyżowskim i ropczycko-sędziszowskim. W wyborach innych niż samorządowe wysoka zmienność widoczna jest w Zagłębiu Dąbrowskim (powiatach: Dąbrowa Górnicza, Sosnowiec). Niska zmienność występuje w całej Polsce wschodniej, z wyjątkiem powiatu hajnowskiego na Podlasiu.

Charakterystyczny jest wykres zmienności poparcia w wyborach innych niż samorządowe. Fakt, że regiony Polski wschodniej charakteryzują się tak niską zmiennością jest bardzo istotny. O ile zmienność w wyborach samorządowych jest rozmieszczona nieco losowo, tak tu zauważamy pewną regułę.



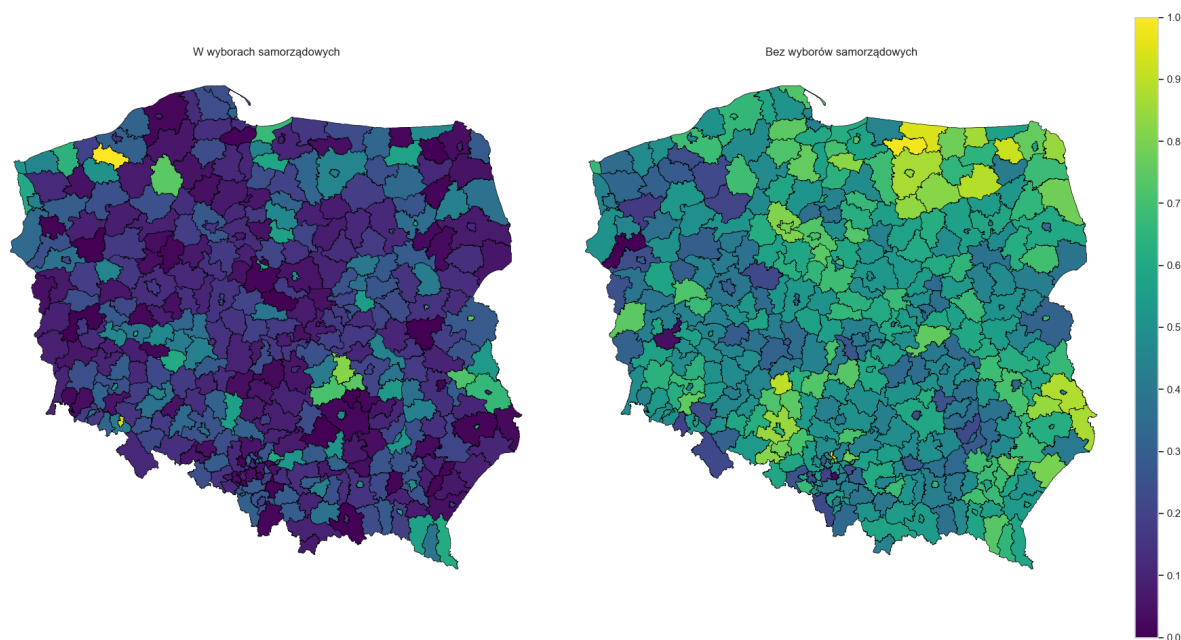
Rysunek 2: Mapy zmienności poparcia dwóch głównych partii

Powyższe mapy zmienności poparcia dwóch głównych partii pokazują coś zgoła odmiennego od map zmienności poparcia wszystkich partii. Po pierwsze możemy zauważyć, że w wyborach samorządowych poparcie sumaryczne Koalicji Obywatelskiej i Prawa i Sprawiedliwości praktycznie się nie zmienia. Inaczej wygląda to wyborach innych niż samorządowe, w których Opolszczyzna jest widocznym regionem, w którym zmienność jest wysoka. Podobnie wysoka zmienność występuje w Zagłębiu Dąbrowskim czy w Bieszczadach.



Rysunek 3: Mapy zmienności poparcia pozostałych sił politycznych

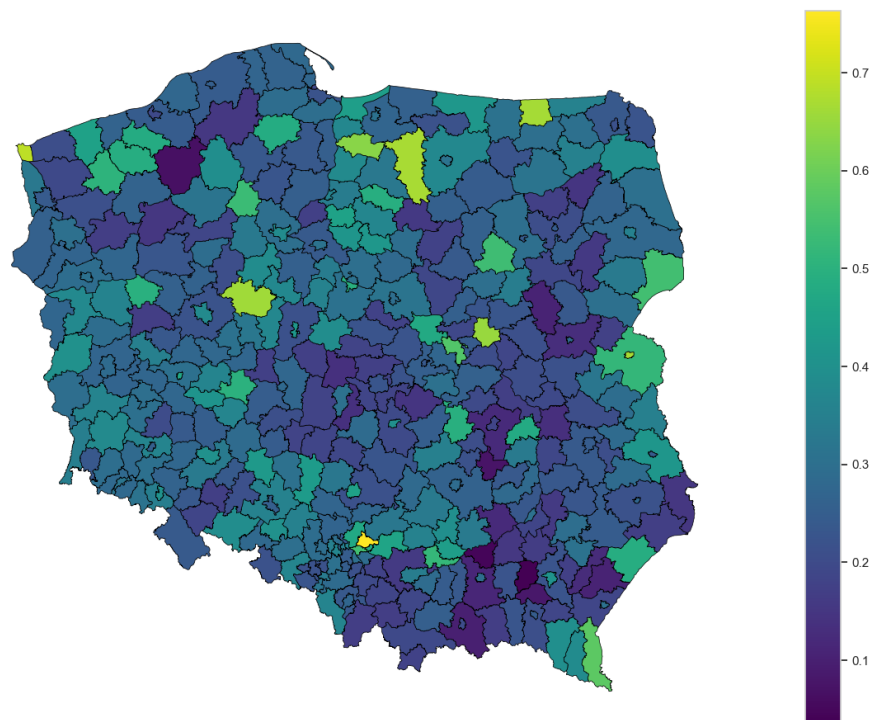
Z tego co jesteśmy w stanie zauważyć, zmienność poparcia reszty partii jest mocno losowa w przypadku wyborów samorządowych, co po raz kolejny dowodzi, że są to wybory lokalne o innej specyfice niż reszta wyborów. Jeśli chodzi o wybory inne niż samorządowe po raz kolejny widać, że w Polsce wschodniej zmienność jest bardzo niska.



Rysunek 4: Mapy zmienności dystansu między Prawem i Sprawiedliwością a Koalicją Obywatelską

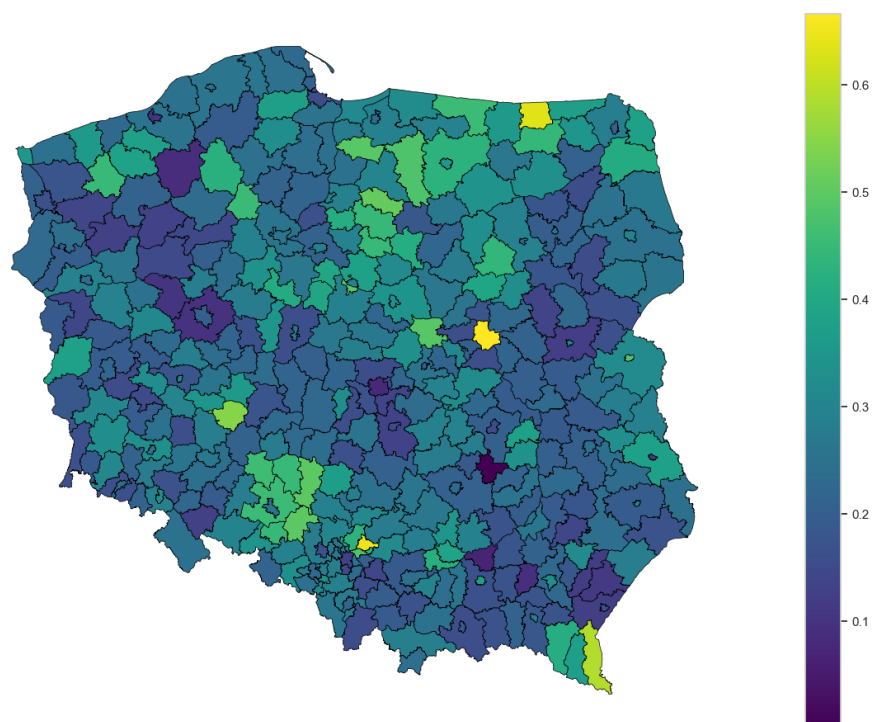
Na powyższych mapach widać, że zmienność różnicy poparcia dwóch głównych sił zmienia się niewiele w wyborach samorządowych, a w wyborach innych niż samorządowe zmienia się znacznie bardziej. Wysoka zmienność w różnicy poparcia występuje w wyborach innych niż samorządowe w Warmii i Mazurach, na Opolszczyźnie, na Chełmszczyźnie i na Kujawach.

4 Mapy skumulowanej zmienności poparcia

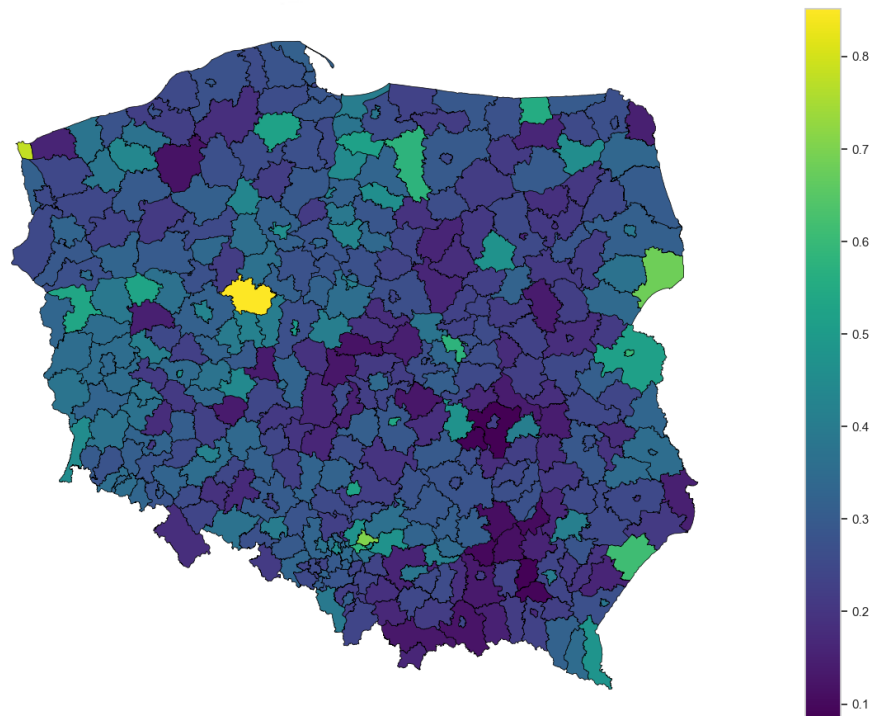


Rysunek 5: Mapa skumulowanej zmienności poparcia

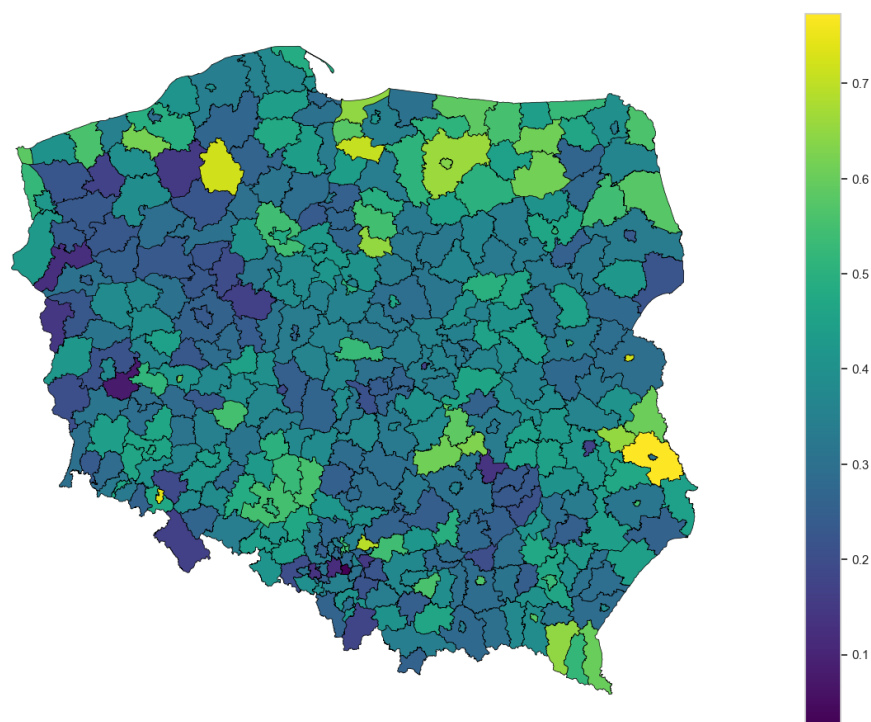
Jak widać, miary zmienności się wypłaszczyły i wystrzyło się to, że miara zmienności poparcia jest w istocie losowa - nie ma konkretnych regionów, które zmieniają swoje preferencje wyborcze. Wgląd w resztę map pozwoli nam poszukać jakichkolwiek zależności między zmiennością poparcia, a konkretnymi regionami.



Rysunek 6: Mapa skumulowanej zmienności poparcia dwóch głównych partii



Rysunek 7: Mapa skumulowanej zmienności poparcia pozostałych partii



Rysunek 8: Mapa skumulowanej zmienności dystansu między Prawem i Sprawiedliwością a Koalicją Obywatelską

5 Badanie korelacji zmienności poparcia z poszczególnymi cechami

Do badania korelacji zebrałem następujące cechy:

- Stopa bezrobocia,
- Ludność w miastach,
- Saldo migracji,
- Dochody własne gminy per capita,
- Miesięczne wynagrodzenia brutto w relacji do średniej krajowej,
- Ludność na aptekę,
- Łóżka w szpitalach ogólnych,
- Współczynnik feminizacji
- Mediana wieku
- Średnia frekwencja
- Szerokość geograficzna
- Długość geograficzna

Zanim przejdziemy do wyników badania, warto omówić dokładnie te cechy:

5.1 Analiza zebranych cech powiatów

5.1.1 Stopa bezrobocia

Miara stopy bezrobocia została pozyskana z danych GUSu. Jest to miara, która pokazuje odsetek osób bezrobotnych. Na potrzeby badania policzyłem średnią stopę bezrobocia w latach 2018-2023.

5.1.2 Ludność w miastach

Miara ludności w miastach, pozyskana z danych GUSu, pokazuje odsetek mieszkańców powiatu, zamieszkujących miasta. Na potrzeby badania policzyłem średnią wartość z lat 2018-2023.

5.1.3 Saldo migracji

Pozyska z danych GUSu miara salda migracji, pokazuje różnicę między liczbą osób, które przybyły do powiatu, a liczbą osób które z niego wyjechały. Na potrzeby badania policzyłem średnią z lat 2018-2023 oraz dokonałem normalizacji min-max tj. przeskalowałem dane do przedziału $\langle 0, 1 \rangle$.

5.1.4 Dochody własne gminy per capita

Miara dochodów własnych gminy per capita, pozyskana z danych GUSu, pokazuje dochody gminy, pochodzące z innych źródeł niż subwencja ogólna i dotacje celowe z budżetu. Im wyższe dochody gminy, tym większe wpływy gminy z m.in podatków od nieruchomości, od środków transportowych, od spadków i darowizn czy od czynności cywilnoprawnych. Miara jest per capita, co oznacza, że liczona jest ona w przeliczeniu na jednego mieszkańca. Na potrzeby badania dane na poziomie gminnym zostały zaagregowane do poziomu powiatu, policzona została średnia z lat 2018-2023, a następnie miara została znormalizowana min-max.

5.1.5 Miesięczne wynagrodzenia brutto w relacji do średniej krajowej

Miara ta pokazuje relację średniego miesięcznego wynagrodzenia brutto pracownika w powiecie w porównaniu do średniej krajowej. Podobnie jak we wszystkich miarach, policzona została średnia z lat 2018-2023. Pozyskana z danych GUSu.

5.1.6 Ludność na aptekę

Oznacza to liczbę mieszkańców przypadających na jedną aptekę w powiecie. Miara jest średnią z lat 2018-2023 oraz została znormalizowana min-max. Pozyskana z danych GUSu.

5.1.7 Łóżka w szpitalach ogólnych

Miara liczby łóżek w szpitalach ogólnych na 10 tys. mieszkańców. Miara jest średnią z lat 2018-2023 oraz została znormalizowana min-max. Pozyskana z danych GUSu.

5.1.8 Współczynnik feminizacji

Współczynnik feminizacji to miara, która pokazuje liczbę kobiet przypadających na 100 mężczyzn w powiecie. Miara jest średnią z lat 2018-2023 oraz została znormalizowana min-max. Pozyskana z danych GUSu.

5.1.9 Mediana wieku

Mediana wieku to miara, która pokazuje wiek, w którym połowa mieszkańców powiatu jest młodsza, a połowa starsza. Miara jest średnią z lat 2018-2023 oraz została znormalizowana min-max. Pozyskana z danych GUSu.

5.1.10 Średnia frekwencja

Średnia frekwencja to miara, która pokazuje średnią frekwencję w wyborach samorządowych i innych niż samorządowe w latach 2018-2023. Miara jest średnią z lat 2018-2024. W tym przypadku miara została policzona na podstawie danych PKW. Jako iloraz liczby kart ważnych do liczby uprawnionych do głosowania.

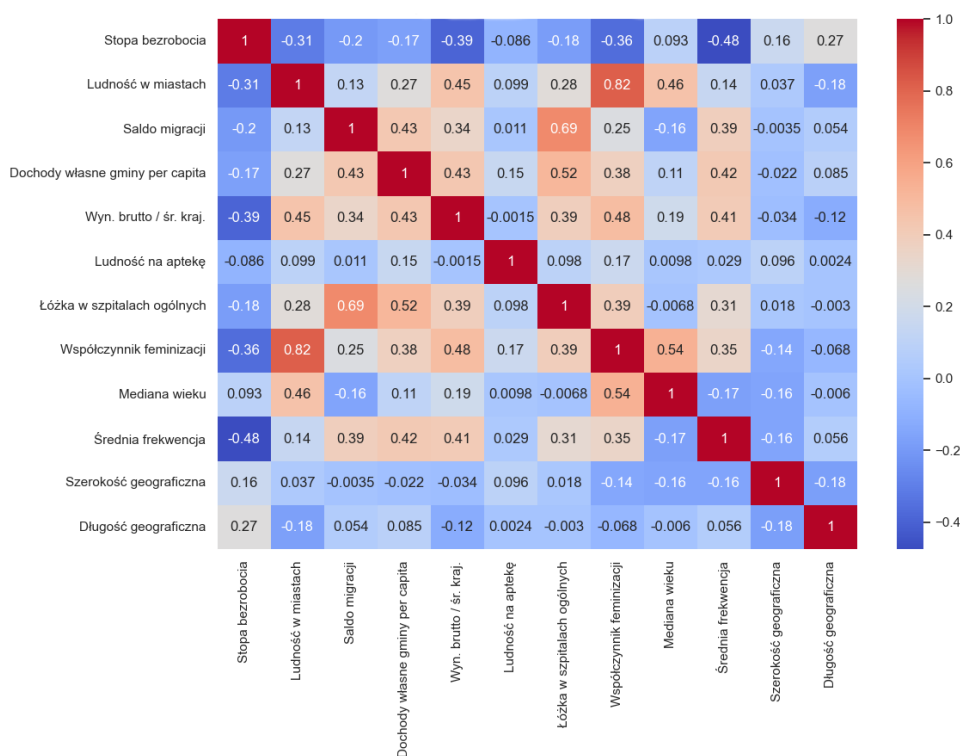
5.1.11 Szerokość geograficzna

Szerokość geograficzna to miara, która pokazuje szerokość geograficzną środka powiatu. Obliczona na podstawie danych geoprzestrzennych.

5.1.12 Długość geograficzna

Długość geograficzna to miara, która pokazuje długość geograficzną środka powiatu. Obliczona na podstawie danych geoprzestrzennych.

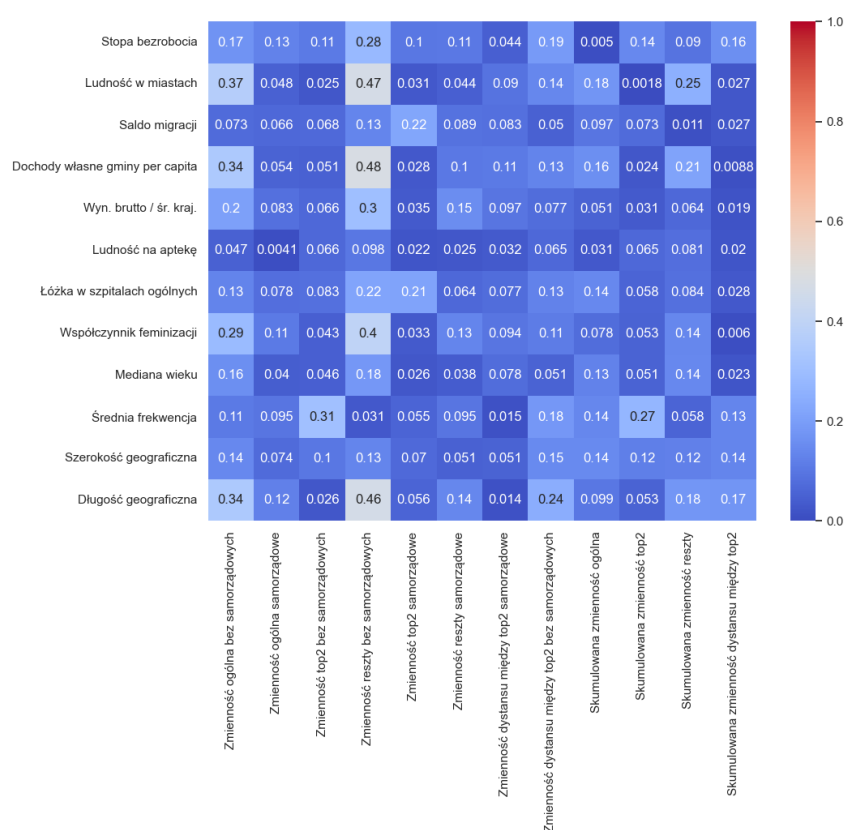
5.2 Analiza korelacji między miarami



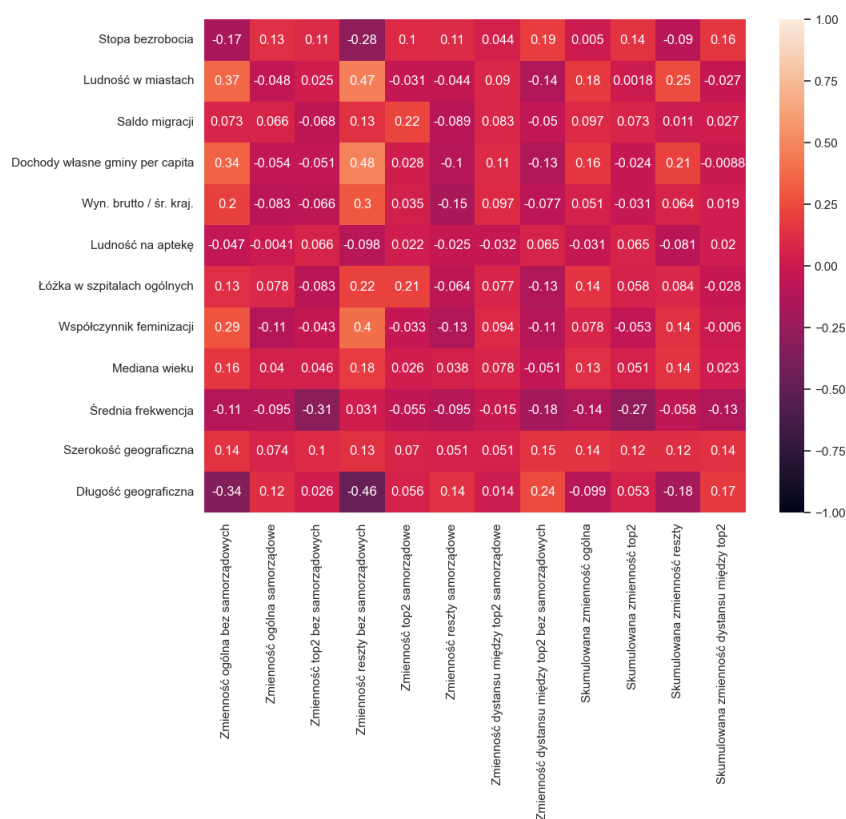
Rysunek 9: Mapa cieplna bezwzględnych wartości korelacji między poszczególnymi miarami

Aby upewnić się, czy cechy nie są ze sobą zbyt mocno skorelowane, sporządziłem macierz korelacji i przedstawiłem ją w postaci mapy cieplnej. Okazuje się, że większość cech jest ze sobą słabo skorelowana, co jest bardzo dobrym prognostykiem, gdyż pozwoli lepiej zbadać miary zmienności. Jest jednak kilka cech, które są mocno skorelowane. Zauważalna jest korelacja między ludnością w miastach a współczynnikiem feminizacji - w większych miastach jest więcej kobiet w przeliczeniu na mężczyzn. Saldo migracji jest średnio skorelowane z liczbą łóżek w szpitalach ogólnych. Wynika z tego, że obie cechy związane są z dużymi miastami. Tym lepiej to widać, gdy spojrzymy na resztę korelacji cech z liczbą łóżek w szpitalach ogólnych. Im bogatszy powiat, tym jest ich więcej - to w miarę sensowne.

5.3 Analiza korelacji poszczególnych cech z miarami zmienności poparcia



Rysunek 10: Mapa cieplna bezwzględnych wartości korelacji między poszczególnymi cechami, a miarami zmienności poparcia



Rysunek 11: Mapa cieplna wartości korelacji między poszczególnymi cechami, a miarami zmienności poparcia

Powyższa mapa cieplna pokazuje, że miara zmienności poparcia jest miarą mocno niezależną i nieskorelowaną w większym stopniu z żadną z wymienionych cech. Rozmaitość miar zmienności poparcia sprawia, że mimo wszystko jesteśmy w stanie dostrzec jakieś zależności. Najciekawszymi miarami pod względem korelacji są z pewnością miary zmienności ogólnej poparcia z wyłączeniem wyborów samorządowych oraz zmienność partii spoza duopolu również bez wyborów samorządowych - raz jeszcze ujawniła się lokalna specyfika wyborów samorządowych, w których zmienność poparcia w powiecie determinują elementy specyficzne dla konkretnych powiatów. Warto zauważyć, że wymienione przeze mnie miary są średnio ujemnie skorelowane z długością geograficzną - im bardziej na wschód, tym mniejsza zmienność i vice versa. Współczynnik feminizacji jest średnio skorelowany z wymienionymi miarami zmienności - to bardzo ciekawa informacja. Warto też zauważyć, że powiaty o wyższym procencie ludności w miastach oraz o wyższych dochodach gminy per capita charakteryzowały się wyższymi wartościami wymienionych miar. Powinniśmy rozumieć to tak, że w ostatnich latach, to właśnie w miastach na prawach powiatu, dysponujących największymi dochodami własnymi, zmieniały się najmocniej wyniki partii spoza duopolu.

6 Podsumowanie

W przeprowadzonym badaniu przeanalizowałem miary zmienności preferencji wyborczych na poziomie powiatów w Polsce na mapach oraz w korelacji z innymi cechami powiatów. Badanie wykazało kilka interesujących zależności. Po pierwsze, wybory samorządowe charakteryzują się znacznie inną specyfiką niż cała reszta wyborów. Gdy na liście komitetu do sejmikowego okręgu pojawia się znane nazwisko, lokalny samorządowiec, znana postać w regionie, wtedy liczba oddanych głosów na ten komitet w regionie może być znacznie wyższa niż realne poparcie dla partii. Zaobserwowaliśmy to na mapach zmienności poparcia wyborczego w wyborach samorządowych, gdzie zmienność poparcia była mocno zróżnicowana w skali całego kraju i ciężko było znaleźć konkretne zależności, które ją determinują. Po drugie, zmienność w wyborach innych niż samorządowe, jest dużo bardziej uzależniona od położenia niż od konkretnych cech powiatu. Jedyne cechy, które jakkolwiek były skorelowane z miarami zmienności poparcia to: procentowy współczynnik ludności w miastach, współczynnik feminizacji $\langle 0, 1 \rangle$, dochody własne gminy i ujemnie skorelowana długość geograficzna. Gdyby złożyć te wszystkie cechy, okazałoby się, że najmniejszą zmiennością wykazują się regiony wschodniej Polski, a największą miasta na prawach powiatu.