



ANALIZA WYNIKÓW WYBORÓW W III RP

*Andrzej Sokołowski, StatSoft Polska Sp. z o.o.; Akademia Ekonomiczna
w Krakowie, Katedra Statystyki*

Jesteśmy w trakcie gorącego sezonu wyborczego. Wybory parlamentarne niemal sąsiadujące w kalendarzu z wyborami prezydenckimi powodują, że od wielu tygodni jesteśmy bombardowani informacjami kampanii wyborczych i komentarzami dziennikarzy. Uodpornienie społeczeństwa na ten atak jest duże, a wiara w możliwość realnego wpływu na rzeczywistość polityczną – mała. Stąd niska frekwencja wyborcza. Zawodowe firmy badania opinii publicznej publikują nie tylko wyniki prostych sondaży, lecz także wyniki w przekrojach grup zawodowych, miejsc zamieszkania, wieku itp. Wiele tygodników drukuje swoiste ściagi dla wyborców, które w schematyczny sposób starają się przedstawić i porównać programy wyborcze partii. Wreszcie dzień sądu nadchodzi i wyborcy udają się do urn (niestety nie wszyscy). Wyniki głosowania często nie pokrywają się z wynikami badań sondażowych, nawet tych przeprowadzanych wśród osób wychodzących z lokali wyborczych. Dobrze udokumentowane jest zjawisko wstydu, polegające na tym, że respondent wstydzi się przyznać ankieterowi, że głosował na ugrupowanie, które w jakiś sposób jest napiętnowane w środkach komunikacji masowej lub opinii publicznej.

W tej prezentacji przedstawione będą wyniki analiz statystycznych wykorzystujący niemal wyłącznie wyniki wyborów wyrażone liczbą (i procentem) głosów. Nie ma w nich znaczenia, co głosi się w kampaniach wyborczych, jakie są wyniki sondaży, co wynika z socjologicznych badań elektoratu, itp. – liczą się tylko surowe liczby, opinie wyrażone „nogami”, czy raczej kartką wyborczą.

Metody statystyczne

Dane statystyczne pochodzą z oficjalnych komunikatów Państwowej Komisji Wyborczej. Podstawowym pomysłem badawczym jest wykorzystanie danych wojewódzkich (mogą to być również dane na poziomie innych jednostek administracyjnych). Tabela danych ma liczbę wierszy równą liczbie województw, zaś liczbę kolumn równą liczbie ugrupowań politycznych lub kandydatów na stanowisko prezydenta. W większości analiz brano pod uwagę tylko te ugrupowania, które zarejestrowały swoje listy we wszystkich województwach. Wartości w takiej tabeli to procenty głosów, jakie uzyskały poszczególne ugrupowania w województwach.



Tak przygotowane dane statystyczne umożliwiają zastosowanie wielu metod analiz wielowymiarowych. Mówiąc językiem tych analiz, badanymi „objektami” mogą być ugrupowania lub kandydaci umieszczone w „przestrzeni” województw. Można też rozpatrywać odwrotną sytuację, gdy obiektami są województwa, a przestrzeń klasyfikacji tworzą ugrupowania lub kandydaci.

We wszystkich analizach socjologicznych podkreśla się przestrzenne zróżnicowanie wyników wyborów. Pewne województwa są tradycyjnymi bastionami pewnych partii. Spostrzeżenie niejednorodności województw, pod względem struktury oddawanych głosów, upoważnia badacza do zastosowania metod taksonomicznych poszukujących bardziej jednorodnych zbiorów obiektów. Popularną metodą jest tu aglomeracyjna procedura Warda. Na wstępie traktuje ona każdy obiekt jako osobną grupę, a potem krok po kroku łączy te obiekty (lub grupy), które powodują najmniejszy przyrost wariancji międzygrupowej. Odległości łączonych podgrup to tzw. odległości aglomeracyjne. Cały proces łączenia jest ilustrowany przez drzewko połączeń zwane dendrogramem. Istnieje ponad trzydzieści różnych metod wskazujących optymalny etap przerywania procesu łączenia, co prowadzi do uzyskania podziału wynikowego. Zaskakująco dobre wyniki daje jednak podział dendrogramu dokonany na podstawie wzrokowej analizy jego struktury – przy czym oczywiście ta identyfikacja liczby grup jest tym lepsza, im większe doświadczenie ma badacz. Po wyliczeniu procentów głosów w grupach (lub średnich procentów w grupach) można interpretować różnice w postawach wyborczych.

Prostym sposobem analizy wyników głosowania jest policzenie współczynników korelacji liniowej pomiędzy wszystkimi partiami. Można się rozsądnie spodziewać, że tam gdzie jest więcej głosów na ugrupowania lewicowe (jakkolwiek tę lewicowość czy prawicowość zdefiniujemy), tam będzie relatywnie mniej na ugrupowania prawicowe. W tym miejscu można do analiz włączyć inne zmienne, na przykład frekwencję wyborczą, stopę bezrobocia, procent ludności miejskiej. Możemy wówczas stwierdzić, które ugrupowania „korzystają” na bezrobociu lub na niskiej frekwencji.

Współczynniki korelacji mogą być traktowane jako miary podobieństwa. Im korelacja jest większa, tym bardziej podobne są odpowiednie ugrupowania polityczne. Ujemne wartości współczynników korelacji świadczą o niepodobieństwie. Można usiłować tak ustawić na płaszczyźnie punkty reprezentujące ugrupowania, aby te skorelowane ujemnie były daleko od siebie, a te skorelowane dodatnio – blisko. Metodą, która zapewnia takie właśnie ułożenie punktów, jest skalowanie wielowymiarowe.

Ciekawym sposobem ilustracji obserwacji wielowymiarowych na płaszczyźnie są tzw. orły. Jest to udoskonalenie znanego wykresu radarowego, w którym osie reprezentujące cechy rozchodzą się równomiernie od środka rysunku. Można rozsądnie zapytać, skąd wiadomo, które cechy powinny być obok których, oraz zakwestionować równość kątów pomiędzy sąsiednimi osiami. To prowadzi nas do koncepcji orłów. Osie są tam wyznaczone położeniem punktów na diagramie otrzymanym ze skalowania wielowymiarowego. Wtedy kąty między osiami są mniej więcej odwrotnie proporcjonalne do współczynników korelacji między nimi.



W niniejszym opracowaniu przedstawiamy tylko wybrane tabele i ilustracje dla niektórych wyborów, jakie miały miejsce w Polsce po roku 1989. Pełna prezentacja pokazana będzie w trakcie seminarium.

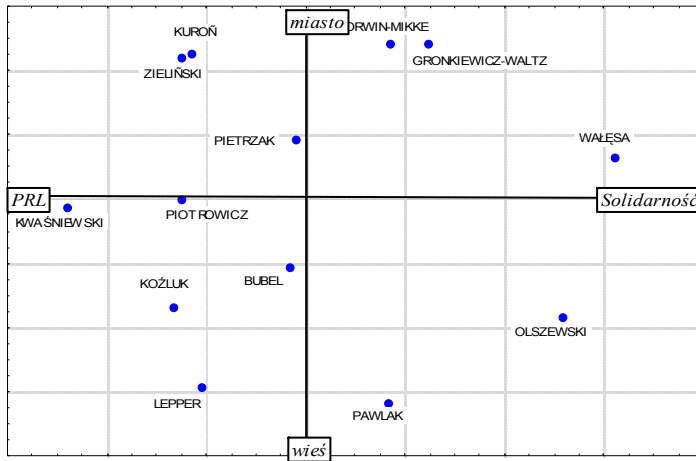
Wybory prezydenckie – 1995

Pierwsza runda tych wyborów miała miejsce w dniu 5 listopada 1995 roku. Poszczególni kandydaci otrzymali następujące procenty głosów ważnych:

Aleksander Kwaśniewski	35,11%
Lech Wałęsa	33,11%
Jacek Kuron	9,22%
Jan Olszewski	6,86%
Władysław Pawlak	4,31%
Tadeusz Zieliński	3,53%
Hanna Gronkiewicz-Waltz	2,76%
Janusz Korwin-Mikke	2,40%
Andrzej Lepper	1,32%
Jan Pietrzak	1,12%
Tadeusz Koźluk	0,15%
Kazimierz Piotrowicz	0,07%
Leszek Bubel	0,04%

Jak wiadomo, drugą rundę wyborów prezydenckich wygrał Aleksander Kwaśniewski. Przy dwóch kandydatach analizy są mało ciekawe, dlatego przyjrzymy się tu tylko wynikom z pierwszej tury. Przypomnijmy, że w Polsce było wówczas 49 województw. Największą korelację ujemną odnotowano – zgodnie z oczekiwaniem – pomiędzy Kwaśniewskim a Wałęsą. Wyniosła ona $-0,91$. Najwyższa korelacja dodatnia wystąpiła pomiędzy Kuronem a Zielińskim ($0,79$). Na wykresie sporządzonym przy wykorzystaniu skalowania wielowymiarowego widać, że najbardziej odległe są od siebie punkty reprezentujące kandydatów, którzy przeszli do drugiej rundy.

Osie wykresu skalowania wielowymiarowego nie posiadają żadnej interpretacji merytorycznej narzucanej przez metodę. W analizie wyników wszystkich wyborów daje się zauważyć, że dla układu punktów mają istotne znaczenie dwa wymiary: poziomy – ideologiczny (lewica-prawica) i pionowy – środowiskowy (miasto-wieś). Interpretacja osi poziomej może wzbudzać pewne kontrowersje w związku z trudnością w zadowalającym zdefiniowaniu biegunów osi ideologicznej. Dlatego na poniższym wykresie zdecydowaliśmy się wprowadzić nazwy odwołujące się do historycznych źródeł kandydatów, ich odniesień ideologicznych i sentymentów elektoratu.

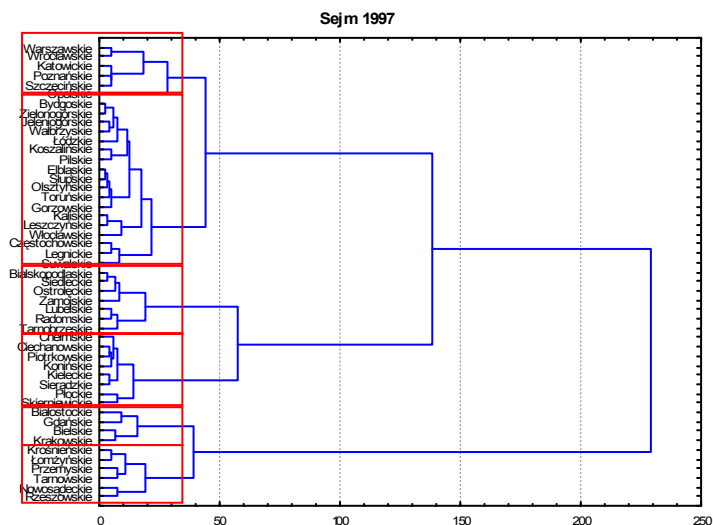


Wybory do Sejmu – 1997

Wybory do Sejmu odbyły się 21 września 1997 roku. Wyniki ogólnopolskie przedstawiały się następująco:

Unia Pracy	4,75%
Blok dla Polski	1,36%
Krajowe Porozumienie Emerytów i Rencistów	1,63%
Unia Wolności	13,38%
Akcja Wyborcza Solidarność	33,86%
Sojusz Lewicy Demokratycznej	27,16%
Polskie Stronnictwo Ludowe	7,31%
Unia Polityki Realnej	2,04%
Ruch Odbudowy Polski	5,56%
Krajowa Partia Emerytów i Rencistów	2,18%
Polska Wspólnota Narodowa	0,07%
Samoobrona	0,08%
Inne	0,62%

W dalszym ciągu nasz kraj był jeszcze podzielony na 49 województw. Po zastosowaniu aglomeracyjnej metody Warda wykres drzewkowy dość wyraźnie sugeruje występowanie aż sześciu grup województw.

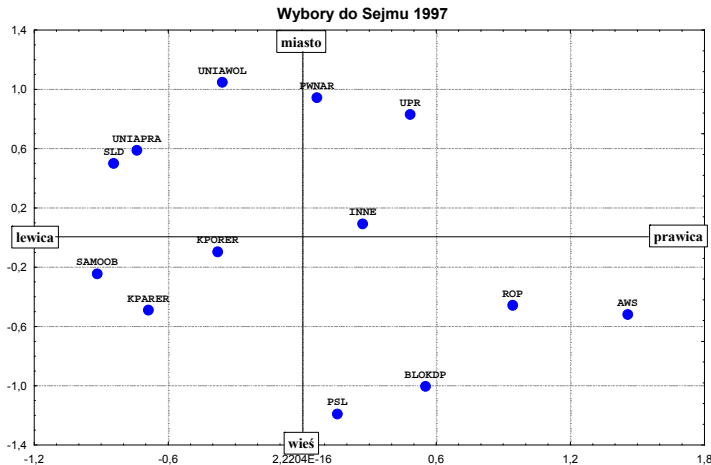


Średnie procenty głosów przedstawia poniższa tabela. Dla ułatwienia interpretacji największe wartości w wierszach wyróżniono drukiem pogrubionym, a najmniejsze – kursywą. Widać, że w grupie pierwszej znalazły się województwa zawierające największe ośrodki miejskie. Unia Wolności była najbardziej popularna właśnie w tej grupie, natomiast zdecydowanie nie głosowano tam na PSL.

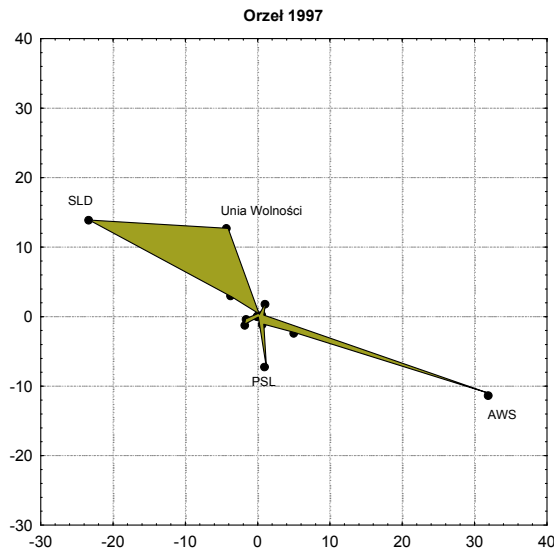
Ugrupowanie	Grupa 1	Grupa 2	Grupa 3	Grupa 4	Grupa 5	Grupa 6
Unia Pracy	5,2	5,8	3,9	5,2	3,7	3,2
Blok dla Polski	<i>0,9</i>	1,2	1,6	1,9	1,3	2,7
KPorozEiR	1,7	1,9	1,8	1,7	<i>1,3</i>	1,5
Unia Wolności	18,0	12,3	<i>7,4</i>	8,9	14,4	7,7
AWS	29,7	27,3	35,2	<i>26,1</i>	45,5	50,9
SLD	28,2	35,1	20,9	29,2	21,4	<i>13,7</i>
PSL	<i>3,6</i>	6,9	16,7	15,9	3,7	10,2
UPR	2,4	1,8	1,9	1,8	1,9	<i>1,6</i>
ROP	4,9	<i>4,5</i>	8,1	6,4	5,1	6,7
KPEiR	2,0	2,7	2,3	2,6	<i>1,6</i>	1,8
KPartiaEiR	0,1	0,0	0,0	0,1	0,0	0,0
PWN	0,1	0,2	0,0	0,1	0,0	0,1
Samoobrona	3,1	0,2	0,1	0,1	0,0	0,0
Inne	5,2	5,8	3,9	5,2	3,7	3,2

Wynik skalowania wielowymiarowego wskazuje na coś, co nastąpiło potem – powstanie koalicji SLD z Unią Pracy. To te dwa punkty są najbliższej, co oznacza, że partie te były

postrzegane przez elektorat jako podobne. Widać też, że Samoobrona zaczęła zdobywać popularność nie tylko na wsi.

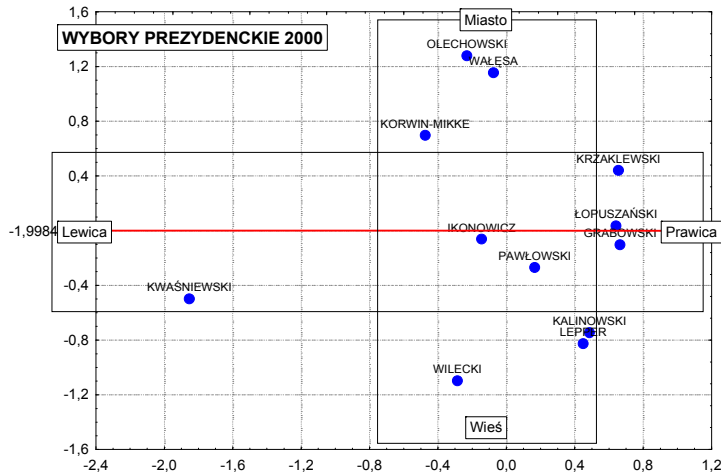


Orzeł wskazuje na dwubiegunowość polskiej sceny politycznej i minimalne znaczenie wielu drobniejszych ugrupowań.



Wybory prezydenckie w 2000 roku

Wybory te wygrał w pierwszej rundzie Aleksander Kwaśniewski, zapewniając sobie prezydenturę na drugą kadencję. Po lewej stronie sceny politycznej nie miał żadnych konkurentów, a głosy zdobywał zarówno w mieście, jak i na wsi.



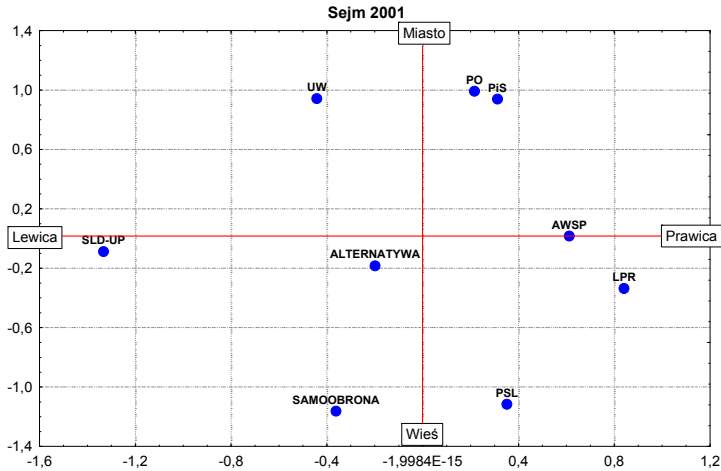
Precyzyjne miejsce punktów na wykresie skalowania wielowymiarowego jest mniej ważne. Ważne są relacje z innymi punktami. Na powyższym wykresie zamiast osi zaznaczono dwa pasy neutralne. Mimo to widać polaryzację kandydatów (poważnych), zarówno ideologiczną jak i środowiskową.

Wybory do Sejmu w 2001 roku

Poprzednie wybory parlamentarne zakończyły się przytłaczającym zwycięstwem SLD. Procentowe wyniki przedstawiały się następująco.

Sojusz Lewicy Demokratycznej - Unia Pracy	41,04%
Platforma Obywatelska	12,68%
Samoobrona Rzeczypospolitej Polskiej	10,20%
Prawo i Sprawiedliwość	9,50%
Polskie Stronictwo Ludowe	8,98%
Liga Polskich Rodzin	7,87%
Akcja Wyborcza Solidarność Prawicy	5,60%
Unia Wolności	3,10%
Alternatywa Ruch Społeczny	0,42%

Rysunek pokazuje dominację koalicji SLD-UP na lewicy, analogiczną do dominacji Aleksandra Kwaśniewskiego w wyborach prezydenckich.



Bardzo ciekawa jest macierz korelacji dla ugrupowań, które weszły do parlamentu.

	SLD_UP%	SAMOOB%	PIS%	PSL%	PO%	LPR%
SLD_UP%	1,00	0,18	-0,64	-0,44	-0,24	-0,74
SAMOOB%	0,18	1,00	-0,54	0,43	-0,66	0,11
PIS%	-0,64	-0,54	1,00	0,03	0,53	0,34
PSL%	-0,44	0,43	0,03	1,00	-0,59	0,63
PO%	-0,24	-0,66	0,53	-0,59	1,00	-0,23
LPR%	-0,74	0,11	0,34	0,63	-0,23	1,00

Ujemne współczynniki korelacji identyfikują pary antagonistów. Najbardziej wyrazista to LPR kontra SLD/UP ze współczynnikiem $-0,74$. Już po zakończeniu tej kadencji można więc stwierdzić, że LPR, „walcząc” z SLD, postępowała zgodnie z wolą swoich wyborców.

Wybory do Sejmu w 2005 roku

Jak doskonale pamiętamy, ostatnie wybory do Sejmu odbyły się w dniu 25 września 2005 roku. Czołowe ugrupowania polityczne uzyskały następujące procenty głosów:

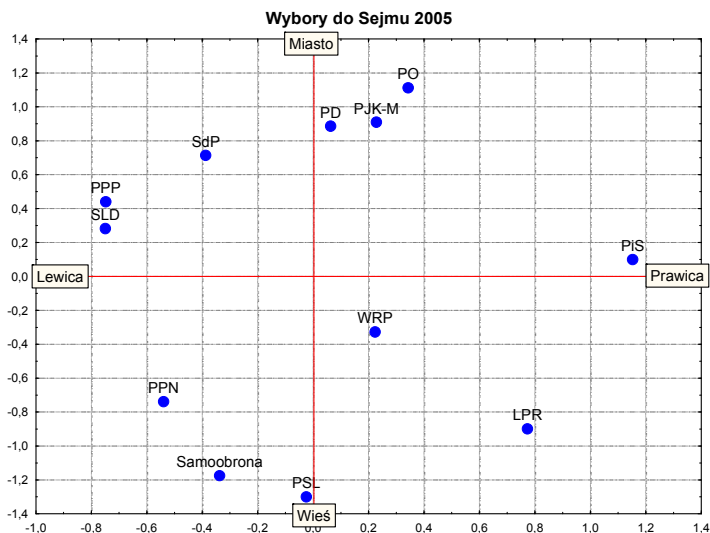
Prawo i Sprawiedliwość	26,99%
Platforma Obywatelska	24,14%
Samoobrona	11,41%
Sojusz Lewicy Demokratycznej	11,31%
Liga Polskich Rodzin	7,97%
Polskie Stronnictwo Ludowe	6,96%



Socjaldemokracja Polska 3,89%

Partia Demokratyczna 2,45%

Do oszacowania macierzy korelacji, a następnie znalezienia konfiguracji punktów w skalowaniu wielowymiarowym wykorzystano dane z 41 okręgów wyborczych. Wzięto pod uwagę te ugrupowania, które zgłosiły listy kandydatów w co najmniej 40 okręgach.



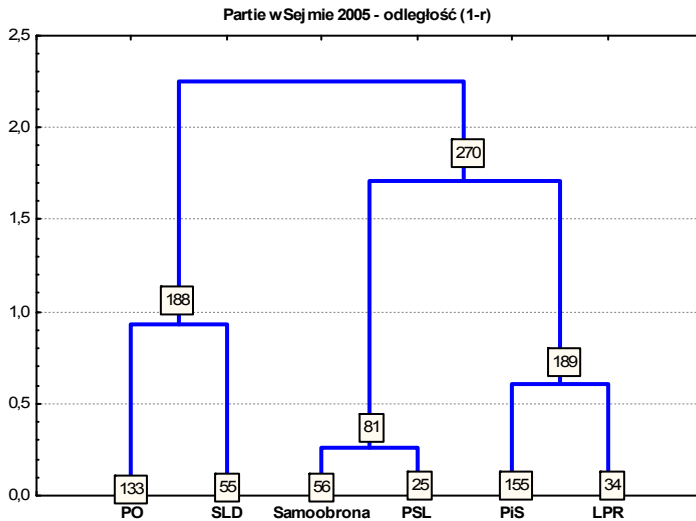
Wydaje się, że w dalszym ciągu ważne są dwa wymiary – ideologiczny i środowiskowy. O ile oś wieś-miasto jest nie do podważenia, to oś ideologiczna znów sprawia kłopoty z nazwaniem jej biegunów. Pewne elementy programu gospodarczego PiS-u są wielce „socjalistyczne”. Ważniejsze wydają się tu elementy światopoglądowe. Zagadnienie to wymaga jednak dalszych rozważań. Warto tu podać jeszcze inną potencjalną interpretację osi pionowej. Byłaby to oś „skorzystania na transformacji ustrojowej” – im wyżej, tym zadowolenie większe.

Ta część tekstu, dotycząca roku 2005, jest pisana w dzień po ogłoszeniu oficjalnych wyników wyborów, po ogłoszeniu kandydatury Kazimierza Marcinkiewicza na premiera, a przed rozmowami pomiędzy ewentualnymi koalicjantami. Macierz korelacji pomiędzy partiami, które weszły do parlamentu, jest dość zaskakująca.

	LPR	PiS	SLD	PO	PSL	Samoobrona
LPR	1,00	0,39	-0,49	-0,59	0,50	0,08
PiS	0,39	1,00	-0,61	0,17	-0,28	-0,59
SLD	-0,49	-0,61	1,00	0,07	-0,16	0,06
PO	-0,59	0,17	0,07	1,00	-0,87	-0,72
PSL	0,50	-0,28	-0,16	-0,87	1,00	0,74
Samoobrona	0,08	-0,59	0,06	-0,72	0,74	1,00

Wynika z niej, że wyborcy raczej widzieliby koalicję PiS z Ligą Polskich Rodzin niż z Platformą Obywatelską. Taka koalicja spajana byłaby wrogością wobec SLD i Samoobrony. Platforma Obywatelska jest partią „oryginalną”, bez istotnej korelacji dodatniej z kimkolwiek. Samoobronę i PSL łączy to samo środowisko – elektorat wiejski.

Na koniec zabawmy się, w ramach statystycznej *political fiction*, w tworzenie koalicji zgodnie z wolą wyborców. Po zastosowaniu metody Warda, w której jako miarę odległości przyjęto ($1 - \text{współczynnik korelacji}$), otrzymujemy następujący dendrogram (nie uwzględnia on Mniejszości Niemieckiej, która zdobyła dwa mandaty, a dla której nie da się wyliczyć współczynników korelacji z innymi partiami).



Liczby na gałęziach pokazują liczbę mandatów. Dopiero dość egzotyczna koalicja wiejsko–„bogo-ojczyzniana” obejmująca PiS, Samoobronę, LPR i PiS jest w stanie przekroczyć próg 50% liczby posłów, który w Sejmie wynosi 230 posłów. Taka koalicja reprezentowałaby 53,33% wyborców, którzy oddali ważne głosy. Czy jest coś, co może łączyć te ugrupowania? Wydaje się, że jest to niechęć do liberalizmu gospodarczego. Brnąc dalej w *political fiction* z dendrogramu, możemy wywnioskować, że najbardziej zrównoważony układ personalny czołowych stanowisk w państwie to: Jarosław Kaczyński jako premier (reprezentant konserwatywnej prawej pary na dendrogramie), Donald Tusk jako prezydent (reprezentant liberalnej pary po lewej stronie) i marszałek Sejmu z wiejskiej pary środkowej. Z uwagi na powagę tego ostatniego urzędu nie odważę się na zaproponowanie przedstawiciela mocniejszego elementu środkowej pary.

A czytelnik tego tekstu będzie już wiedział, co nam przyniosła *political reality*, czyli *polityczna rzeczywistość*.