

# Ocena zaufania w stosunku do Parlamentu Europejskiego w Niemczech w 2014 roku. Analiza skupień

*Aleksandra Ciepluch, Anna Rokicka, Krzysztof Wojdalski*

*2017-01-24*

## Wstęp

Zaufanie do instytucji publicznych jest jednym z filarów pozwalających funkcjonować instytucjom demokratycznym. Duży jego poziom daje silny mandat na przeprowadzanie reform i efektywną pracę struktur rządowych. Kryzys finansowy, i w konsekwencji konieczne wsparcie finansowe dla najbardziej zadłużonych krajów, a także rosnący problem z napływem imigrantów i terroryzmem - każdy z tych czynników powoduje, że Unia Europejska zmagą się z ogromnym kryzysem zaufania. Odsetek deklarujących wiarę w instytucję Unii Europejskiej spadł z 57% w 2007 roku do 31 % w roku 2012 i od tego utrzymuje się obecnie na niemal niezmiennym poziomie. Celem naszego projektu jest pogrupowanie respondentów na grupy w zależności od poziomu zaufania do Parlamentu Europejskiego, stopnia zainteresowania polityką, wykształcenia i wieku. W tym celu wykorzystamy analizę skupień.

## Analiza skupień - teoria

Analiza skupień ma na celu podział danych na grupy (klastry/skupienia), zawierające elementy o podobnych cechach. Wykorzystywany algorytm działa w taki sposób, aby jednostki w obrębie danej grupy były jak najbardziej do siebie podobne, ale poszczególne wyodrębnione grupy cechowały się jak najmniejszym podobieństwem. W ramach analizy skupień wyróżnia się dwa główne typy:

- Metoda hierarchiczna - polega na stworzeniu kolejnych podziałów zbioru, zaczynając od takiego, w którym każdy element tworzy samodzielne skupienie, poprzez połączenie obiektów leżących najbliżej siebie, aż do uzyskania jednego skupienia. Problem stanowi wybór metryki, mierzącej odległości między poszczególnymi skupieniami. Do najczęściej używanych należą:
- Metoda najbliższego sąsiedztwa, w której to odległość między dwoma skupieniami określona jest przez odległość między najbliższymi obiektami (sąsiadami) należącymi do różnych skupień; Metoda najdalejszego sąsiedztwa, tutaj odległość, w przeciwieństwie do metody najbliższych sąsiadów równa jest odległości między najdalej położonymi obiektami należącymi do różnych skupień;
- Metoda średnich połączeń, gdzie odległość między skupieniami równa się średniej arytmetycznej wyznaczonej ze wszystkich odległości obiektów należących do dwóch różnych skupień;
- Metoda środków ciężkości - odległość wyznaczona jest poprzez odległość między środkami ciężkości skupień
- Metoda ważonych środków ciężkości (mediany), metoda podobna do ważonych środków ciężkości z tą różnicą, że w tym przypadku wprowadzamy wagi, odpowiadające wielkością skupień.
- Metoda niehierarchiczna - w tym przypadku, w przeciwieństwie do metody hierarchicznej żadna z grup na które został podzielony zbiór danych, nie jest podgrupą innej grupy z tego zbioru.

Najbardziej popularną odmianą niehierarchicznej analizy skupień jest metoda k-średnich, w której zbiór jest dzielony na z góry ustaloną liczbę skupień. Poszczególne elementy są przyporządkowywane za pomocą algorytmu, który ma na celu pogrupowanie elementów o jak największym podobieństwie do jednej grupy, przy jednoczesnej maksymalizacji różnic pomiędzy odrębnymi skupieniami. Głównym założeniem analizy skupień jest losowy wybór próby oraz jak najmniejsza korelacja pomiędzy zmiennymi niezależnymi.

## Baza danych

W naszej analizie korzystaliśmy z bazy danych Europejskiego Badania Społecznego, przeprowadzonego w 2014 roku. Zbiór zawierał odpowiedzi 3045 respondentów z Niemiec na pytania dotyczące życia społecznego. Wywiady zostały przeprowadzone w przeważającej liczbie przypadków metodą kontaktu bezpośredniego. W dalszej analizie wyselekcjonowaliśmy te zmienne, które uznaliśmy za najbardziej adekwatne w kontekście naszego badania. Ponadto z badania wykluczaliśmy obserwacje zawierające niepełne odpowiedzi. Ostatecznie zbiór składał się z 2804 obserwacji i to właśnie one stanowiły podstawę naszego badania. Wybrane zmienne na podstawie literatury przedmiotu i wstępnej inspekcji danych są następujące: \* interested (polintr) - zmienna opisująca poziom zainteresowania polityką, \* trust (trstprl) - zmienna opisująca zaufania respondentów do Parlamentu Europejskiego przyjmuje wartości od 0 do 10) \* education (edulvlb) - zmienna opisująca poziom edukacji respondenta zgodnie z międzynarodową standardową klasyfikacją edukacji, \* wiek (agea) - zmienna opisująca wiek respondenta. Dla ułatwienia badania pierwotne nazwy (w nawiasach) zostały zmienione na bardziej intuicyjne.

## Wybór metody badawczej

Do naszego badania wybraliśmy metodę niehierarchiczną. Decyzja była uwarunkowana specyfiką problemu. Po pierwsze nie znaleźliśmy podstaw do hierarchizacji naszych grup. Po drugie - baza charakteryzuje się dużą liczbą obserwacji, co utrudniłoby analizę wszystkich zmiennych w przypadku metody hierarchicznej. Jednym z jej wymogów jest policzenie macierzy zmiennych o wymiarze  $n \times n$ , gdzie  $n$  jest liczbą obserwacji. Złożoność obliczeniowa algorytmu jest bardzo duża w przypadku tak dużej liczby zmiennych.

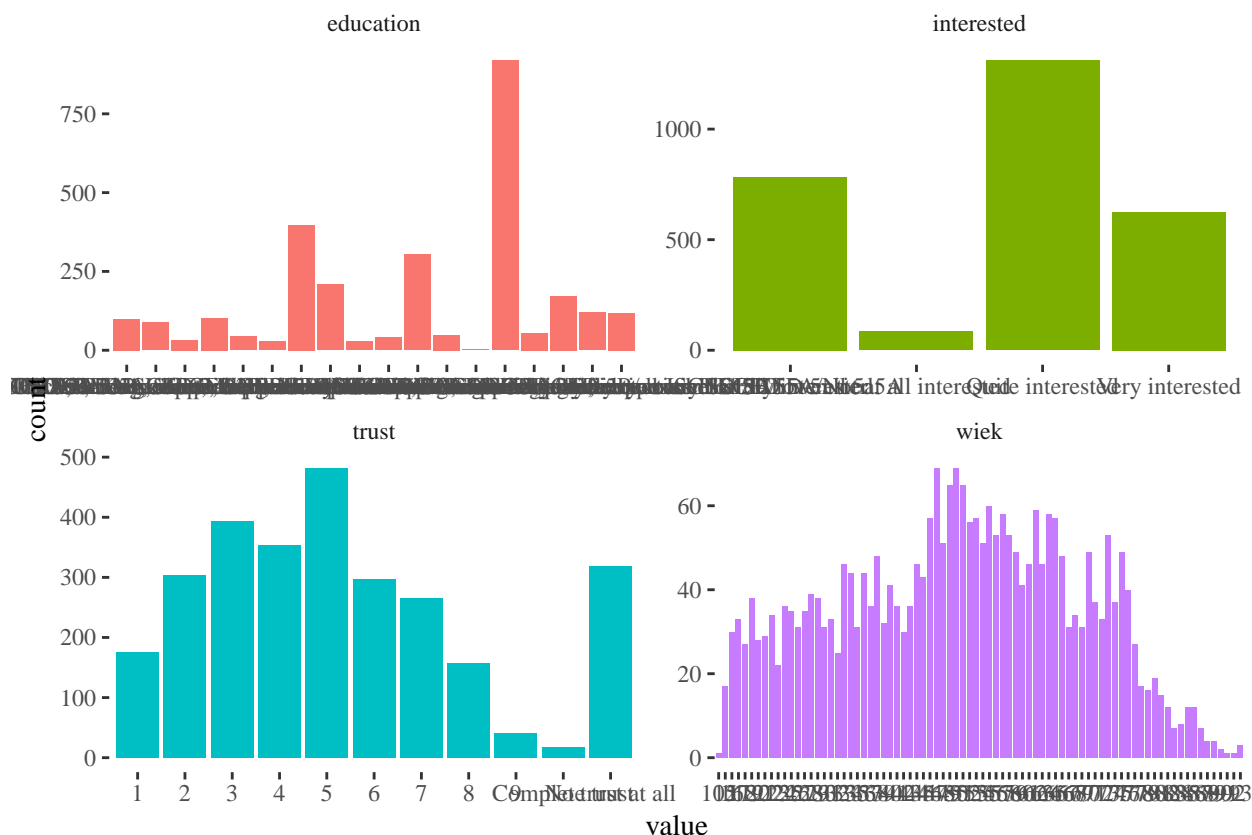
## Hipotezy badawcze

W celu odpowiedzi na nasze pytanie badawcze postawiliśmy kilka hipotez, które brzmią następująco: \* Osoby starsze ufają Parlamentowi mniej - takie stwierdzenie jest motywowane faktem, że ludzie starsi są gorzej wyedukowani i nie w pełni rozumieją benefity wynikające z partycypacji we wspólnym rynku i integracji europejskiej \* Osoby z wyższym wykształceniem ufają Parlamentowi bardziej - takie stwierdzenie jest \* Osoby zainteresowane polityką ufają Parlamentowi najbardziej - zakładamy, że Unia Europejska jest tworem, który w sumie pomaga obywatelom krajów członkowskich w osiągnięciu dobrobytu. W związku z tym ci, którzy zdają sobie sprawę z dobrych stron demokratycznego reżimu i konkurencji na rynku będą ufać bardziej instytucji Parlamentu Europejskiego

## Interpretacja wyników badania

### Wstępna analiza danych

W początkowej fazie badania za pomocą formuł des, oraz sum dokonaliśmy podstawowej analizy zmiennych, poprzez przedstawienie ich podstawowych charakterystyk, takich jak średnia, odchylenie standardowe, obserwacje minimalne i maksymalne (Tabela 1), analizę histogramu dla zmiennej wiek (Wykres 1), oraz macierzy korelacji, obliczonej za pomocą współczynnika Spearmana (Tabela 2).



Standaryzacja zmiennych Ponieważ zmienne są przedstawione w różnych skalach musimy dokonać ich standaryzacji, dzięki czemu w bazie tworzone są nowe zestandaryzowane zmienne odpowiednio `wiek_K`, `interested_K`, `trust_K`, `education_K`. Zdecydowaliśmy się aby w naszym badaniu skorzystać z niehierarchicznej metody k-średnich. Liczbę skupień wyznaczyliśmy za pomocą statystyki Celińskiego i Harabasz., która posiada tym większe wartości im większe zróżnicowanie pomiędzy skupieniami. Wyznaczenie liczby skupień Dla naszych zmiennych jej wartości przedstawiają się następująco:

Liczba skupień Wartość Statystyki Celińskiego- Harabasz 1 - 2 828,61 3 882,81 4 820,19 5 742,82 6 733,65 7 719,91 8 670,63 9 665,76 10 654,32 11 629,85 12 617,9 13 605,04 14 577,04 15 543,04 16 548,99

Ponieważ szukana statystyka osiąga największą wartość dla 3 skupień, oznacza to, że po podzieleniu zbioru danych na dwie grupy osiągniemy największe zróżnicowanie pomiędzy skupieniami. Następnie w celu scharakteryzowania powstałych skupień, przedstawiliśmy charakteryzujące je statystyki opisowe dla poszczególnych zmiennych w postaci tabelarycznej: Charakterystyki poszczególnych skupień

W pierwszej grupie znaleźli się respondenci charakteryzujący się stosunkowo wysokim zaufaniem do Parlamentu Europejskiego, oraz wysokim zainteresowaniem do polityki. Z drugiej strony są to respondenci młodzi o niższym poziomie edukacji, niż w pozostałych grupach.

W drugim skupieniu znajdują się respondenci o niskim zaufaniu do Parlamentu Europejskiego oraz nieco niższym zainteresowaniu polityką. Są to także respondenci lepiej wykształceni oraz starsi, o średniej wieku powyżej 60 lat.

Respondenci z ostatniej grupy charakteryzują się przeciętnym zaufaniem do Parlamentu i zdecydowanie najniższym spośród wszystkich grup poziomie zainteresowania polityką. Są to też osoby najlepiej wykształcone i w średnim wieku. W celu potwierdzenia czy nasze wyniki nie powstały w sposób przypadkowy przeprowadziliśmy dodatkowy test, grupujący zbiór danych za pomocą mediany, która jest bardziej odporna na obserwacje odstające. Wyniki testu przedstawiliśmy w Tabelach 7-9.

Potwierdziło to nasze wnioski, że najbardziej Parlamentowi Europejskiemu ufają osoby najmłodsze o sto-

sunkowo niskim poziomie edukacji i wysokim zainteresowaniu polityką. Z drugiej strony najniższe zaufanie wykazują osoby najstarsze o przeciętnym zaufaniu polityką i średnim poziomie edukacji. Osoby w wieku średnim o zdecydowanie najwyższym poziomie edukacji ufają Parlamentowi w stopniu niewiele niższym jak osoby młode, jednak wykazują zdecydowanie mniejsze zainteresowanie polityką. Weryfikacja hipotez Na podstawie przeprowadzonego zweryfikowano postawione wcześniej hipotezy: Osoby starsze ufają Parlamentowi mniej. Osoby z wyższym wykształceniem ufają Parlamentowi bardziej. Osoby zainteresowane polityką ufają Parlamentowi najbardziej. W przypadku wszystkich hipotez możemy stwierdzić, że nie ma podstaw do ich odrzucenia. Wnioski W trakcie badania natrafiłszy na wiele problemów związanych z danymi. Baza charakteryzuje się dużą rozbieżnością danych, co utrudniło nieco nasze badanie. Niepokojące są niskie korelacje pomiędzy zmiennymi. Intuicja podpowiada, że zmienne wiek i edukacja powinny być mocno skorelowane, jednak w naszych danych ta korelacja jest dosyć niska, nawet wprowadzenie tylko 3 poziomów w przypadku badanych zmiennych nie dało pożądanych skutków i korelacje nie uległy poprawie. Analiza segmentacji wyróżniła 3 niewykluczające się grupy, które

Bibliografia //musimy tu coś dodać pod nasze nowe badanie i uporządkować

E. Uslaner, Trust and Corruption, w: Corruption and the New Institutional Economics, red. M. Taube, J. Lambsdorf, M. Schramm, Routledge, London 2004.

Arnold Christine, Sapir V. Eliyahu, Zapryanova Galina, „ Trust in the institutions of the European Union: A cross-country examination”, 2012, dostępny w internecie : <http://eiop.or.at/eiop/pdf/2012-008.pdf> Załączniki