Raport z Analizy Wielowymiarowej

ANALIZA DYSKRYMINACJI

//beznadziejna strona, ale nie to jest tu najważniejsze! xD

1. **Cel projektu**

Celem projektu jest znalezienie takich zmiennych dyskryminujących, które pozwolą wyjaśnić różnice pomiędzy grupami respondentów oraz pozwolą na przewidzenie przynależności grupowej respondentów. Tematem badania jest zaufanie w stosunku do Parlamentu Europejskiego w Niemczech w 2014 roku. W celu zbadania skonkretyzowania badania postawiono pytanie badawcze brzmiące następująco:

*Czy za pomocą wybranych czynników demograficznych i psychograficznych możemy wyestymować poziom zaufania do Parlamentu Europejskiego ?*

1. **Opis bazy danych**

Wykorzystana baza danych została pobrana ze strony:

Baza zawiera odpowiedzi 3045 respondentów z Niemiec, którzy odpowiadają na określone pytania z zakresu życia społecznego. Wszystkie dane zostały zebrane w 2014 roku. Baza pierwotnie zawierała 601 zmiennych.

1. **Wykorzystywane zmienne**

Na podstawie literatury podmiotu oraz przeprowadzonych „research” wybrano 9 zmiennych, co do których uznano, że dobrze różnicują grupy badanych. Zmiennymi tymi są:

* *happy*-zmienna opisująca poziom szczęścia, na skali //usunęłabym tą zmienną bo nie ma wg mnie uzasadnienia w podziale na grupy
* *interested*- zmienna opisująca poziom zainteresowania polityką, na skali
* *influence* – zmienna opisująca czy system polityczny pozwala ludziom wpływać na politykę, na skali
* *health* – zmienna opisująca poziom zdrowia, na skali
* *gender* – zmienna określająca płeć respondenta
* *education* – zmienna określająca poziom edukacji, na skali 1-5
* *voted* – zmienna określająca, czy respondent głosował w ostatnich wyborach, na skali 0-1
* *petition* – zmienna określająca, czy respondent podpisał jakąś petycję w przeciągu ostatnich 12 miesiecy, miezrona na skali
* *left* – zmienna określająca umiejscowienie na skali lewica-prawica, w skali left(0)-right(10)

1. **Teoretyczne podstawy analizy determinacji**

**//TEORIA, ktoś coś? :P**

1. **Zmienna definiująca**

Zmienna definiująca podział populacji na grupy to *trust*, która określa poziom zaufania do Parlamentu Europejskiego (zmienna przyjmuje wartości od 0 do 10, z badania wykluczyliśmy 121 respondentów , którzy nie odpowiedzieli na to pytanie). Ze względu na dużą liczbę poziomów zmiennej, postanowiono podzielić je na 3 grupy wskazujące na poziom zaufania. Odpowiedzi respondentów podzielono następująco :

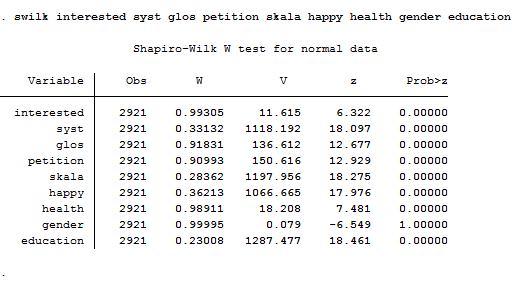
* 7-10 Respondent ma zaufanie wysokie
* 4-6 Respondent ma zaufanie przeciętne
* 0-3 Respondent ma zaufanie niskie

1. **Interpretacja wyników**

// wyniki są dla wszystkich wybranych zmiennych, możecie sobie to zmienić w kodzie, albo dać znać, to ja sprawdzę wyniki, bo pewnie Krzysiek nie ma staty, a ja nie chcę się bawić z kodem w R, bo mamy gotowy kod do staty

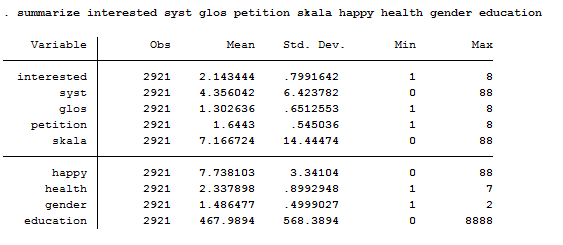
Korzystając z kodów dostępnych z zajęć oraz dostępnej literatury przeprowadzano badanie. Na wstępie zbadano statystyki opisowe badanych zmiennych oraz normalność rozkładów.

Testowanie normalności shapiro wilk

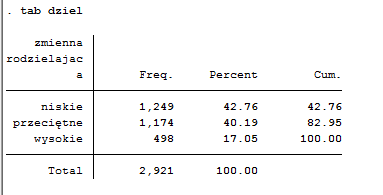


W niektórych przypadkach test wskazuje na to, że rozkład badanej zmiennej odbiega od rozkładu normalnego, jednak po analizie histogramów oraz wykresów kwantyl-kwantyl stwierdzono, że dane mogą pochodzić z rozkładu normalnego.

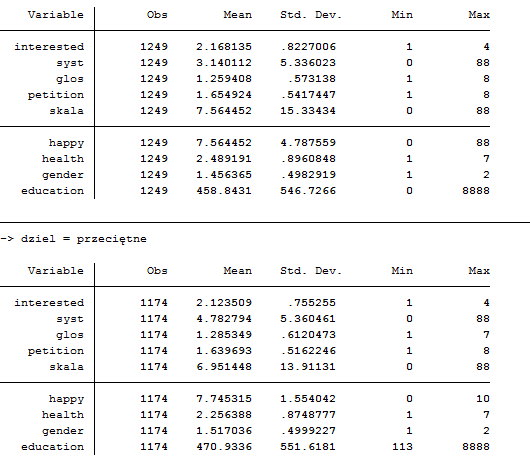
statystyki opisowe badanych zmiennych

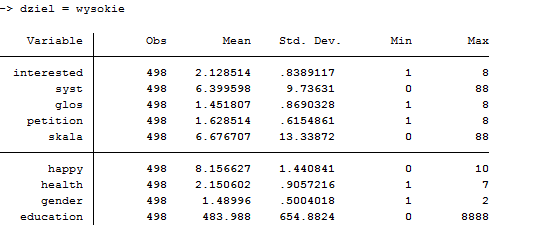


W kolejnym kroku zbadano rozkład zmiennej zależnej, która rozdziela naszą próbę na 3 grupy wg poziomu zaufania. Wyniki testu przedstawiono w tabeli:

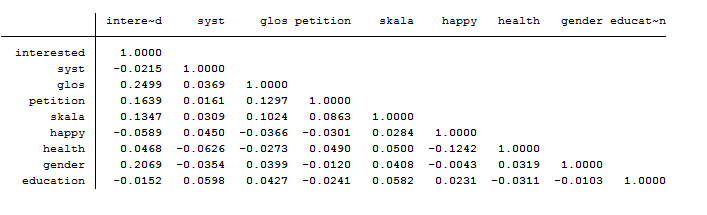


Po przeprowadzeniu podstawowych badań postanowiono dopasować model. //tutaj trzeba się rozpisać i pomyśleć co by tu naściemniać żeby było ładnie no i jaki model…bo nie wiem za bardzo…

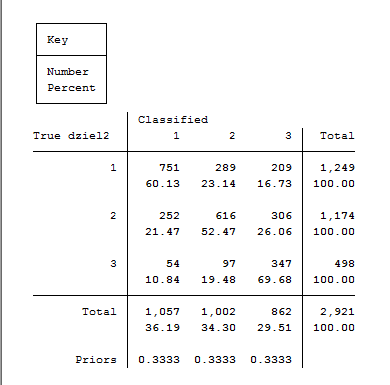




Korelacja między zmiennymi



Wybieramy rozkład a priori i badamy poprawność // znowu trzeba wymyślić coś ładnego



1. **Wnioski**
2. **Bibliografia**

* Gatnar E., Nieparametryczna metoda dyskryminacji i regresji, PWN, Warszawa 2001
* Morrison D.F., Wielowymiarowa analiza statystyczna, PWN, Warszawa 1990
* Klecka, W. R. (1981). Discriminant analysis. Beverly Hills: Sage Publications.
* Lachenbruch, P. A. (1975). Discriminant analysis. New York: Hafner Press.
* 8 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