



## PODSTAWOWE ANALIZY I WIZUALIZACJA Z WYKORZYSTANIEM MAP W STATISTICA

*Krzysztof Suwada, StatSoft Polska Sp. z o.o.*

### Wstęp

Wiele różnych analiz dotyczy danych opisujących wielkości charakterystyczne bądź silnie związane z pewnym obszarem geograficznym, np. dane o zarobkach czy bezrobociu. Celem prezentacji jest przedstawienie analizy danych tego typu i metod wizualizacji jej wyników.

Dane wykorzystywane w opisywanym przykładzie pochodzą z bazy danych GUS, a dotyczą migracji ludności między województwami w 2006 roku.

### Analiza

Rozpoczniemy od otwarcia arkusza z danymi w programie *STATISTICA*.

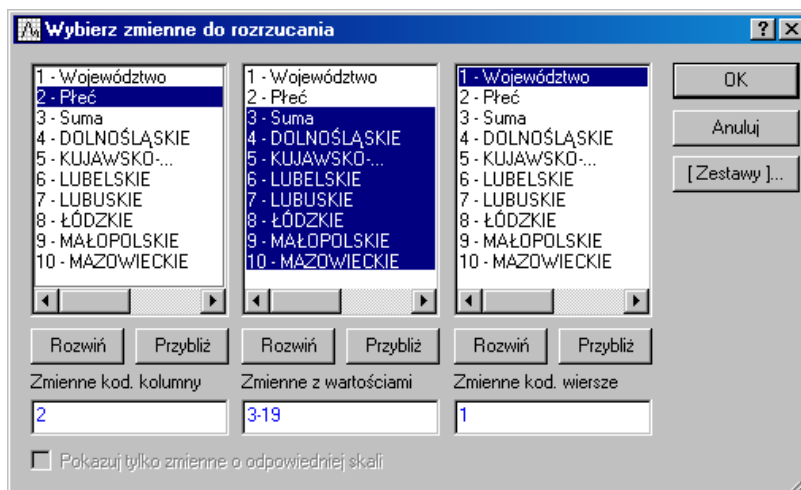
	1 Województwo	2 Płeć	3 DOLNOŚLĄSKIE	4 KUJAWSKO-POMORSKIE
4	LUBUSKIE	Mężczyzna	522	57
5	ŁÓDZKIE	Mężczyzna	196	144
6	MAŁOPOLSKIE	Mężczyzna	235	73
7	MAZOWIECKIE	Mężczyzna	648	715
8	OPOLSKIE	Mężczyzna	351	21
9	PODKARPACKIE	Mężczyzna	102	31
10	PODLASKIE	Mężczyzna	39	30
11	POMORSKIE	Mężczyzna	165	692
12	ŚLĄSKIE	Mężczyzna	354	155
13	ŚWIĘTOKRZYSKIE	Mężczyzna	81	24
14	WARMIŃSKO - MAZURSKIE	Mężczyzna	56	186
15	WIELKOPOLSKIE	Mężczyzna	649	671
16	ZACHODNIOPOMORSKIE	Mężczyzna	243	246
17	DOLNOŚLĄSKIE	Kobieta	15609	152
18	KUJAWSKO-POMORSKIE	Kobieta	140	11624
19	LUBELSKIE	Kobieta	89	48
20	LUBUSKIE	Kobieta	544	69
21	ŁÓDZKIE	Kobieta	244	172
22	MAŁOPOLSKIE	Kobieta	259	89
23	MAZOWIECKIE	Kobieta	699	759
24	OPOLSKIE	Kobieta	445	31

Rys. 1. Arkusz danych.



Jedno z pytań, jakie może postawić badacz, brzmi: „Czy kobiety różnią się od mężczyzn, jeżeli chodzi o skłonność do migracji?”. W tym celu dodajmy dodatkową zmienną zawierającą sumę liczby osób, które zmieniły miejsce zamieszkania w każdym z województw (suma zmiennych od 4 do 19). Ze względu na brak normalności rozkładu w celu porównania kobiet i mężczyzn posłużymy się testami nieparametrycznymi dla prób zależnych. W tym celu musimy przekształcić arkusz do odpowiedniego formatu.

Przekształcenie łatwo wykonać, korzystając z transformacji zawartej w menu „Dane|Rozrzucić po zmiennych...”.

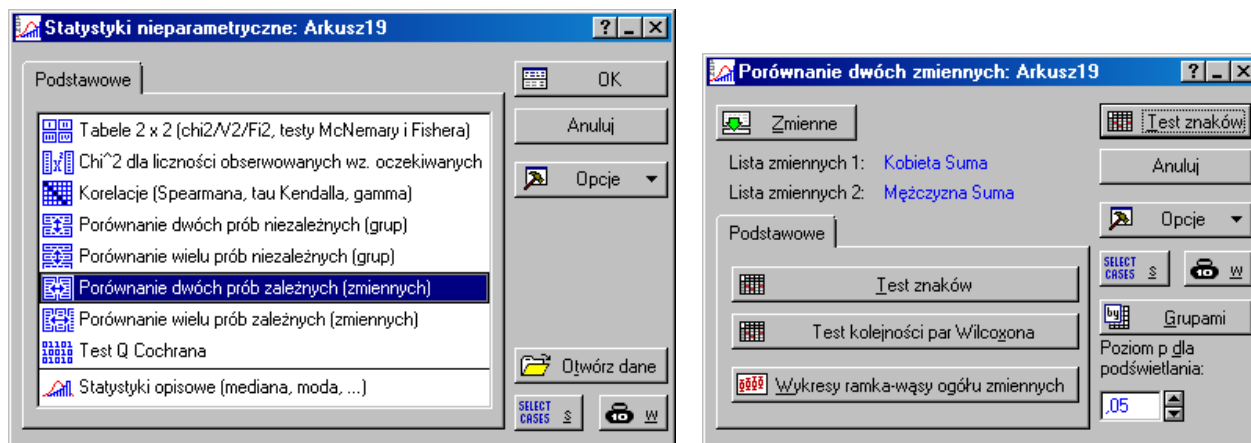


Rys. 2. Okno wyboru zmiennych.

	1 Województwo	2 Kobieta Suma	3 Kobieta DOLNOŚLĄSKIE	4 Kobieta KUJAWSKO- POMORSKIE	5 Kobieta LUBELSKIE	6 Kobieta LUBUSKIE
1	DOLNOŚLĄSKIE	20147	15609	152	231	585
2	KUJAWSKO-PO	14374	140	11624	55	68
3	LUBELSKIE	13182	89	48	11090	45
4	LUBUSKIE	7523	544	69	49	5533
5	ŁÓDZKIE	13870	244	172	134	62
6	MAŁOPOLSKIE	18298	259	89	299	74
7	MAZOWIECKIE	43437	699	759	2906	219
8	OPOLSKIE	5935	445	31	41	40
9	PODKARPACKI	11038	118	27	338	17
10	PODLASKIE	7492	46	32	82	15
11	POMORSKIE	18043	171	805	176	78
12	ŚLĄSKIE	24304	400	190	282	125
13	ŚWIĘTOKRZYS	6896	103	24	104	10
14	WARMIŃSKO -	9767	58	175	73	31
15	WIELKOPOLSK	25020	834	792	141	682
16	ZACHODNIOPO	12076	252	225	122	321

Rys. 3. Arkusz po przekształceniu.

W tym momencie dane są już gotowe do analizy. Wybieramy z menu „Statystyka|Statystyki nieparametryczne”, a następnie „Porównanie dwóch prób zależnych”, a jako zmienne wybieramy sumy dla kobiet i mężczyzn.



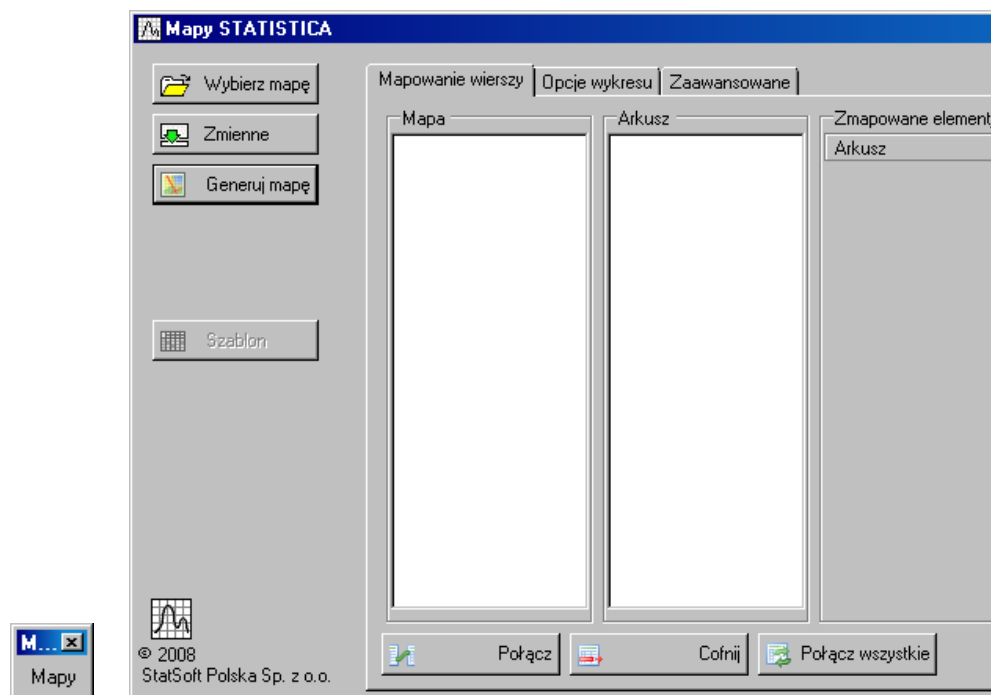
Rys. 4. Porównanie dwóch prób zależnych.

Wyniki analizy wskazują, że istnieje istotna różnica między kobietami, a mężczyznami jeżeli chodzi o gotowość do zmiany miejsca zamieszkania.

Zaznaczone wyniki są istotne z $p < ,05000$			
Liczba	Procent	Z	poziom p
Niewiąz.	$v < V$		
16	0,00	3,750000	0,000177

Rys. 5. Wynik testu znaków.

Otrzymane wyniki możemy w programie *STATISTICA* przedstawić na wiele sposobów, jednym z nich jest wizualizacja na mapach administracyjnych Polski. W tym celu ze strony [www.StatSoft.pl](http://www.StatSoft.pl) należy pobrać i zainstalować dostępny dla użytkowników wersji 8 *STATISTICA* dodatek *Mapy*. W programie pojawi się wtedy dodatkowy pasek narzędzi.

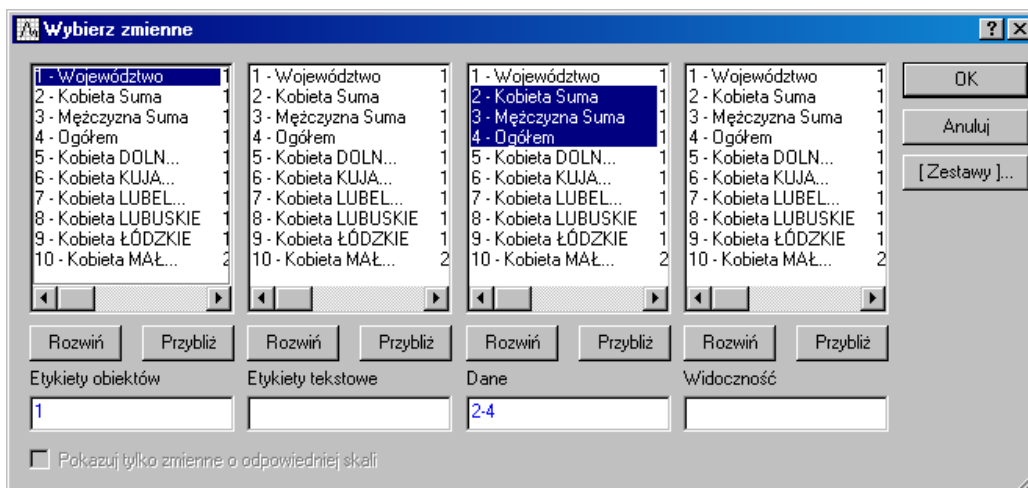


Rys. 6. Program *Mapy*.



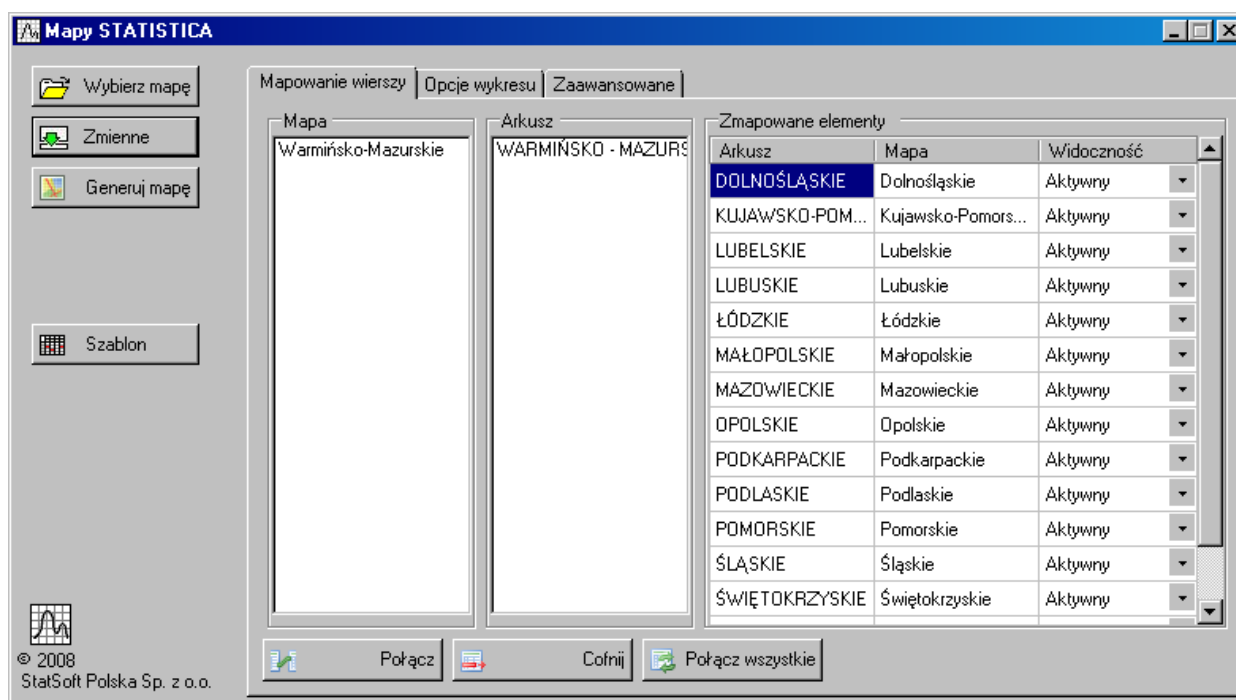
Dodajmy jeszcze jedną zmienną będącą sumą kobiet i mężczyzn, którzy zmienili miejsce zamieszkania – nazwijmy ją „Ogółem”.

Uruchamiamy teraz program *Mapy*, następnie wybieramy zmienne identyfikujące województwa i dane do wizualizacji – „ogółem”, „kobieta suma” i „mężczyzna suma”.



Rys. 7. Okienko wyboru zmiennych.

Po kliknięciu OK program automatycznie połączy województwa z elementami arkusza, natomiast elementy, które nie zostały połączone automatycznie, należy połączyć ręcznie.

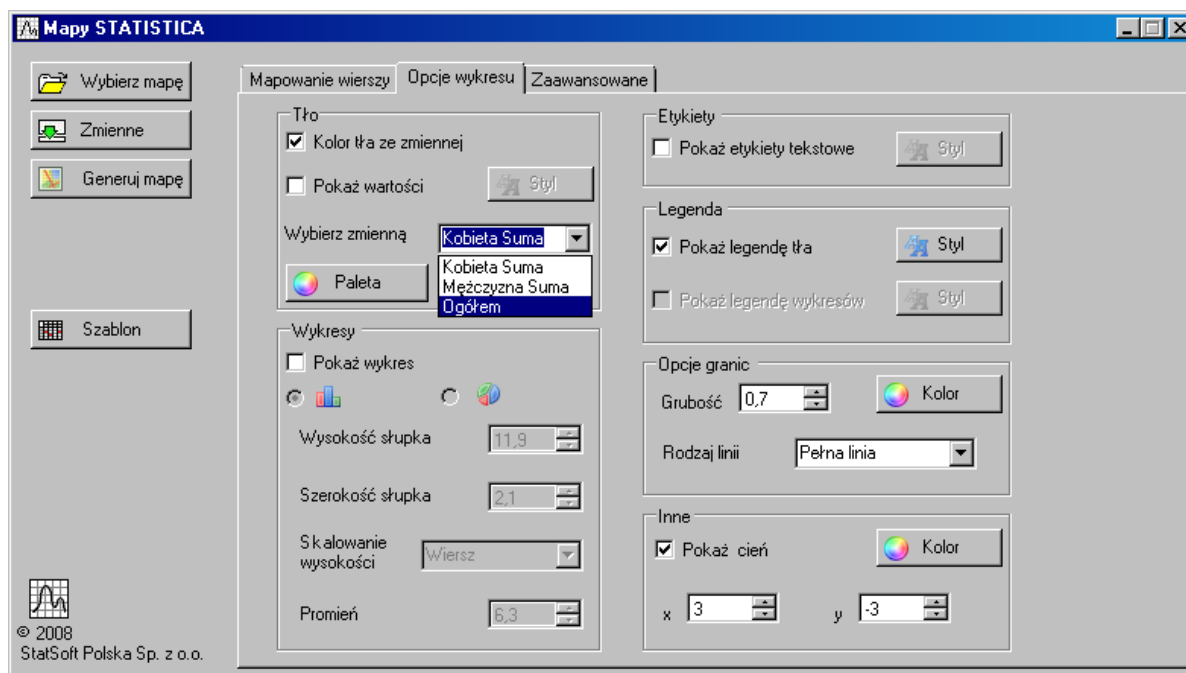


Rys. 8. Okno główne programu *Mapy*.

Opcje wyglądu mapy, takie jak: paleta kolorów, style i kolor granic i inne, można zmieniać na zakładkach „Opcje wykresu” i „Zaawansowane”. W tej przykładowej analizie



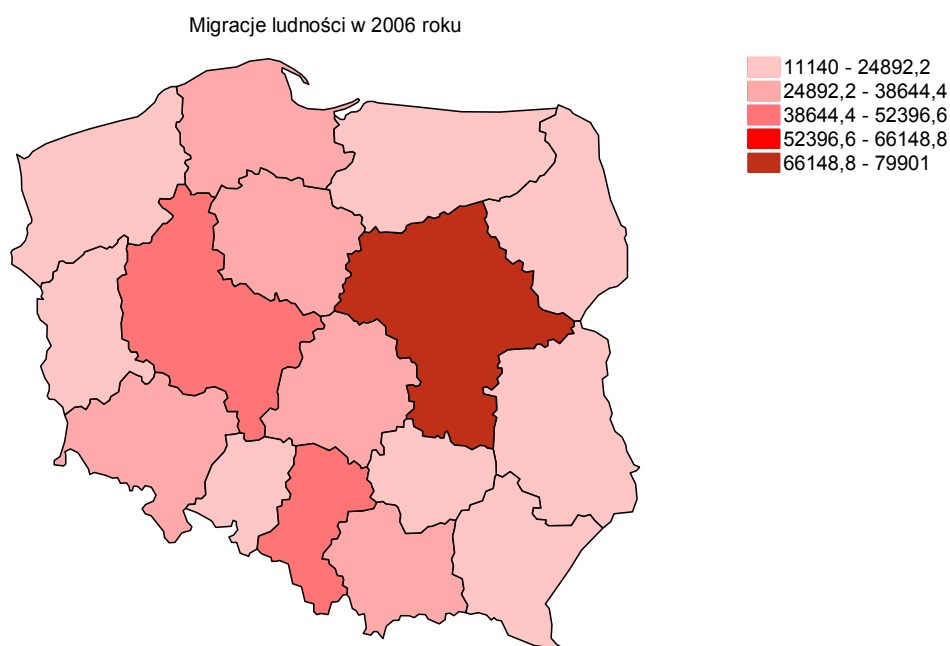
wybraliśmy więcej niż jedną zmienną zawierającą dane, więc wykresy słupkowe i kołowe są dostępne.



Rys. 9. Program *Mapy*, zakładka z opcjami.

Pierwszym krokiem będzie wyłączenie wyświetlania wykresów, następnie należy włączyć opcję „Kolor tła ze zmiennej” i jako zmienną z listy rozwijalnej wskazać „Ogółem”.

Po kliknięciu przycisku „Generuj mapę” otrzymamy pierwszą mapę wygenerowaną przez program.



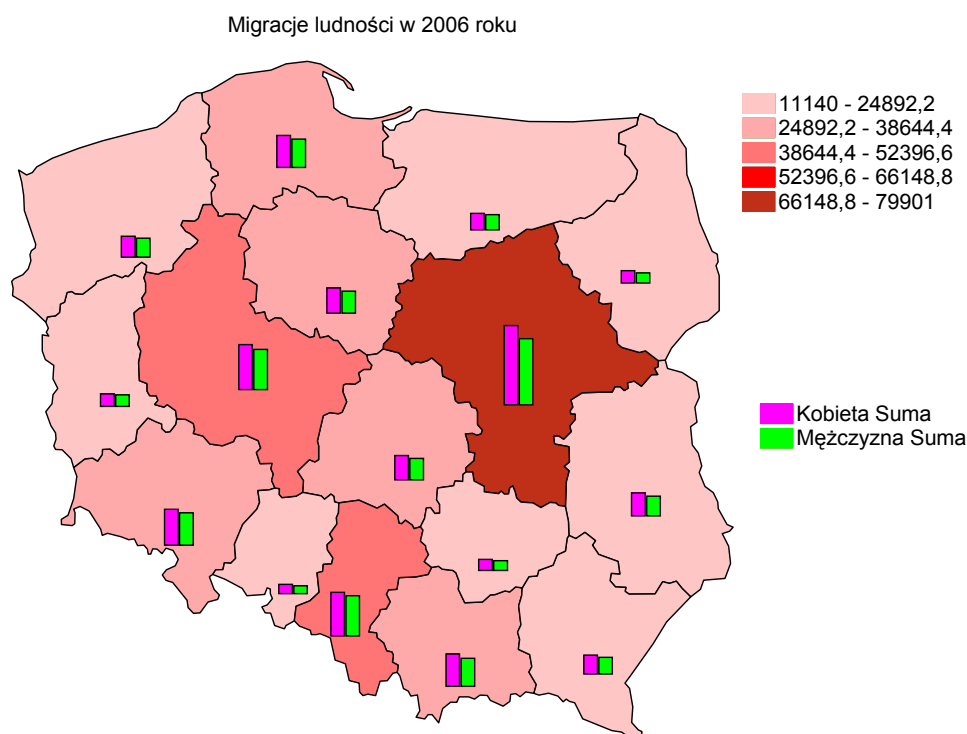
Rys. 10. Pierwsza wygenerowana mapa.



Przydatną informacją jest także udział kobiet i mężczyzn w migracji w podziale na poszczególne województwa. Aby był sens porównywać je nie tylko między sobą w obrębie województwa, ale także między województwami, w opcjach skalowania należy wybrać „globalnie”.



Rys. 11. Ustawienia skalowania.



Rys. 12. Druga wygenerowana mapa.

Wszystkie wyniki można bardzo łatwo umieścić w skoroszycie, a następnie przenieść je do dokumentu MS Word, dzięki temu w szybki i prosty sposób możemy utworzyć kompletny raport z analizy.

## Podsumowanie

Jak potwierdziły testy statystyczne, kobiety i mężczyźni różnią się istotnie, jeżeli chodzi o skłonność do zmiany miejsca zamieszkania. Jak widać na wykresie mieszkańcy województwa mazowieckiego byli najbardziej skłonni do migracji i także w tym województwie dużo więcej kobiet zmieniło miejsce zamieszkania.