Boris Kožuh

# ANALIZA DANYCH W BADANIACH

Kraków 2006

Recenzje:

prof. dr hab. Larisa Głazyrina
prof. dr hab. Tatiana Łopatik

Projekt okładki: Joanna Sroka

Redakcja, adiustacja i łamanie: Saša Kožuh

Korekta: Katarzyna Siwiec



Na zlecenie Krakowskiej Szkoły Wyższej im. Andrzeja Frycza Modrzewskiego
http://www.ksw.edu.pl

ISBN 83-89823-56-X

Wydawca: Krakowskie Towarzystwo Edukacyjne sp. z.o.o. - Oficyna Wydawnicza AFM, Kraków 2006

Druk i oprawa COPY2000, 31-072 Kraków, ul. Wielopole 16

# Spis treści

## JAK OBLICZYĆ:



## JAK WYKONAĆ:

# Wprowadzenie

Drogi Czytelniku!

Przed Tobą książka, która jest niezwykle prostą instrukcją sprawnego poruszania się po labiryncie programów statystycznych. Współczesne programy komputerowe pozwalają bowiem opracowywać dane zgromadzone w badaniach, w sposób łatwy i szybki.

Empiryczne badania edukacyjne przebiegają często w postaci sondażu diagnostycznego. Oznacza to stosunkowo duże populacje lub próby i jednocześnie wielką ilość badanych zmiennych. Zebrany materiał empiryczny przekracza zazwyczaj możliwości ręcznego opracowania statystycznego. Jeszcze dwadzieścia lat temu opracowaniami zajmowały się centra obliczeniowe na uniwersytetach. Wówczas opracowania statystyczne były najbardziej czasochłonnym etapem badań empirycznych. Rozwój małych komputerów umożliwił każdemu badaczowi samodzielne opracowywanie danych. Dziś gromadzenie danych zabiera więcej czasu niż ich opracowanie statystyczne.

Dane można opracowywać za pomocą kilku programów. Najobszernejszym z nich jest SPSS. Podobnie jak i inne programy, on również ewoluuje i jest nieustannie doskonalony. Można przewidzieć, iż program ten będzie dobrze służył jeszcze przez dłuższy czas. Niniejsza książka opisuje zasady użycia programu SPSS 12.0 PL. Pragnę jednak nadmienić, że dobre zrozumienie podstaw korzystania z wersji SPSS 12.0 PL umożliwi czytelnikowi szybkie dostosowanie się do każdej kolejnej wersji SPSS dla Windows.

Do korzystania z programu SPSS niezbędna jest podstawa, którą jest wiedza ze statystyki. W tym celu zalecam moją książkę pod tytułem »Statystyka«, wydaną przez Regionalny Ośrodek Doskonalenia Nauczycieli w Częstochowie w 2005 roku. W książce tej czytelnik znajdzie teoretyczne podstawy wszystkich metod statystycznych omawianych w niniejszym podręczniku.

Pomimo iż w poniższej książce została opisana polska wersja SPSS, do omawianego tekstu dodano nazwy wszystkich poleceń i podprogramów w języku angielskim. Zatem do korzystania z wersji angielskiej nie jest potrzebne dobre opanowanie języka angielskiego.

Książka posiada rozszerzone marginesy. W bardziej teoretycznych rozdziałach na marginesach umieszczono podstawowe pojęcia, wskazówki oraz streszczenia. Zadbano również o to, aby wyeksponować w nich najważniejsze i najistotniejsze wyjaśnienia. W rozdziałach, w których opisuję procedury uruchamiania rozmaitych funkcji, na marginesach zostały umieszczone instrukcje dotyczące korzystania z poszczególnych opcji i poleceń programu SPSS w języku angielskim (nazwy poleceń, opcji oraz przycisków). W ten sposób czytelnik, który posiada tylko angielską wersję programu będzie mógł, korzystając z tekstu i jednocześnie z marginesu, łatwo opanować wszystkie procedury stosowania programu SPSS. Dodatkowo w aneksach zostały umieszczone rysunki najistotniejszych okien programu SPSS w wersji angielskiej.

Szczególnie pragnę polecić dodatek do spisu treści (strona 6), który został opracowany w taki sposób, aby ułatwić znalezienie w książce stron, na których zademonstrowano sposoby obliczania i wykonywania zadań statystycznych.

Mam nadzieję, że lektura poniższej książki przekona każdego, iż statystyka może być prosta, łatwa i zrozumiała, a przede wszystkim niezwykle użyteczna. Przygotowałem ją z myślą głównie o ambitnych pedagogach, ciekawych świata, badających otaczającą rzeczywistość edukacyjną i opracowujących zgromadzony materiał badawczy. Przyjemnej lektury i powodzenia oraz satysfakcji z samodzielnie dokonanych analiz statystycznych.

*Pragnę złożyć serdeczne podziękowania firmie SPSS Polska z siedzibą w Krakowie za udostępnienie polskiej wersji systemu SPSS oraz za wyrażenie zgody na opublikowanie ekranów z tego systemu.*

Autor

# Program SPSS

# Program SPSS

Praca z programem SPSS przebiega w następujących fazach:

### 1. Przygotowanie danych do opracowania

Dane, które chcemy opracować statystycznie za pomocą komputera, najpierw przeglądamy, sortujemy i przygotowujemy (na papierze).

### 2. Wprowadzanie danych do komputera

Przygotowane dane należy wprowadzić do komputera. Program SPSS ma swój podprogram do wprowadzania danych, tak samo można wprowadzić dane za pomocą innych programów (Excel, Word, Access itd.). Jeżeli czytelnik często korzysta z programu Word – jego użycie jest najlepszym rozwiązaniem. Dane wprowadzone za pomocą programu Word należy „importować" do SPSS przed właściwym opracowaniem. Jest to bardzo proste, więc nie ma potrzeby uczyć się procedur wprowadzania za pomocą podprogramu SPSS.

### 3. Opracowanie danych

Wybierając różne polecenia w SPSS uruchamia się wybrane opracowania statystyczne. Po uruchomieniu poleceń otwiera się okno dialogowe. W tym oknie ustawia się różne opcje opracowania.

### 4. Przegląd wyników

Kilka sekund po uruchomieniu końcowego polecenia otwiera się okno Edytor raportów, w którym można zobaczyć wszystkie wyniki. Okno to umożliwia przegląd wyników, zachowanie ich na dysku oraz wydruk.

### 5. Dodatkowe porządkowanie danych

Często po przeglądzie pierwszych wyników powstaje potrzeba powrotu do danych wyjściowych po to, aby je dodatkowo uporządkować (pogrupować, selekcjonować itd.).

### 6. Ponowne opracowanie i przegląd wyników

Po dodatkowym uporządkowaniu ponownie uruchamia się polecenia dla opracowań. Po końcowym opracowaniu tabele z wynikami należy zapisać i zachować na dysku.

### 7. Wykorzystanie wyników

Po opracowaniu wyników i ich przeglądzie wszystkie rezultaty należy przenieść do innych programów i do tekstu, w którym będą analizowane, prezentowane oraz interpretowane (do pracy magisterskiej, dyplomowej, do raportu badań, artykułu itd.).

# 2

# Porządkowanie danych

*Z tego rozdziału nauczysz się:*

- *porządkować dane jakościowe*
- *porządkować dane ilościowe*
- *tworzyć zestawy tabelaryczne*
- *tworzyć szereg szczegółowy prosty*
- *tworzyć szereg rozdzielczy z przedziałami klasowymi*
- *przygotowywać dane do opracowania komputerowego*
- *tworzyć spis zmiennych*

# Wstęp

porządkowanie danych

przygotowanie danych

Przed opracowaniem statystycznym należy uporządkować dane. Czynność ta ułatwia opracowywanie danych. Od czasu, kiedy pojawiły się komputery, procedury porządkowania danych zmieniły się w sposób zasadniczy. Przed "ręcznym" opracowaniem danych należy je uporządkować, natomiast do opracowania za pomocą komputera wystarczy je jedynie przygotować. Aby lepiej zrozumieć wyniki opracowania statystycznego należy zapoznać się z całkowitym postępowaniem: od przygotowania danych do otrzymania wyników, aby znać też drogę do ich otrzymania.

# Porządkowanie danych nominalnych i porządkowych

zestawy tabelaryczne

Dane porządkuje się tworząc zestawy tabelaryczne. Dla każdej kategorii zmiennej należy policzyć ilość jednostek (liczebność). Liczebność zwykle wyrażana jest w postaci liczb absolutnych (np. ilu respondentów wybrało pojedyncze odpowiedzi na pytanie w ankiecie) i liczb procentowych (ile procent respondentów wybrało daną odpowiedź).

pierwsza analiza

Tak uporządkowane dane nadają się do kolejnych opracowań. Już samo uporządkowanie danych umożliwia pierwszą analizę.

***Przykład***

Tabela1. Tabela grupy pracowników według płci

tabela zawiera jedną zmienną

| płeć | liczebność | procent liczebności |
|---|---|---|
| mężczyźni | 25 | 43,1 |
| kobiety | 33 | 56,9 |
| razem | 58 | 100,0 |

Jeżeli zmienna jest porządkową i ma więcej niż dwa stopnie, można dodać kolumnę liczebności skumulowanych F (lub F%). Skumulowana liczebność jest sumą jednostek (lub %) do pewnej kategorii. Dla każdej kategorii należy zsumować wszystkie liczebności (zarówno jej, jak i niższych kategorii). Ma to sens jedynie wtedy, gdy jest to niezbędne.

liczebności skumulowane

Wyniki w statystyce na ogół zaokrągla się do dwóch miejsc po przecinku. Istnieje jednak pewien wyjątek: procenty w tabelach zaokrągla się do jednego miejsca po przecinku. Nie dotyczy to procentów, z których oblicza się kolejne wyniki. Tam znowu obowiązuje prawo dwóch miejsc po przecinku.

dwa miejsca po przecinku

W taki sposób przebiega porządkowanie danych dla wszystkich nominalnych i porządkowych zmiennych w badaniu. Te tabele pokazują stan w badanej populacji. Często oprócz stanu po pojedynczych zmiennych bada się także związki między zmiennymi. Do takich celów należy sporządzić tabele, które zawierają więcej zmiennych. Są to tabele krzyżowe (korelacyjne). Najczęściej taka tabela zawiera dwie zmienne, ponieważ tabela z trzema zmiennymi jest już bardzo skomplikowaną i trudną do odczytania (nawet dla specjalistów).

stan w badanej populacji

tabele krzyżowe

Poniżej zaprezentowano przykład tabeli dwóch zmiennych: stopień wykształcenia i poglądy dotyczące badanego zjawiska.

Tabela 2. Tabela korelacyjna grupy pracowników według wykształcenia i poglądów

| | zgadzam się | niezdecydowany | nie zgadzam się | razem |
|---|---|---|---|---|
| średnie | 12 | 6 | 28 | 46 |
| licencjackie | 5 | 3 | 7 | 15 |
| magisterskie | 11 | 2 | 4 | 17 |
| razem | 28 | 11 | 39 | 78 |

tabela z dwiema zmiennymi

Z tej tabeli niełatwo zrozumieć, jaki jest związek zmiennych. Należy wyliczyć jeszcze procenty we wszystkich komórkach tabeli. Procenty można obliczyć trzema sposobami, otrzymując w ten sposób trzy odmienne tabele (nr 3, 4, 5). Zostały one umieszczone na kolejnych stronach.

procenty w komórkach tabeli

procenty obliczone według stopnia wykształcenia

W tabeli nr 3 procenty zostały obliczone według stopnia wykształcenia, w tabeli nr 4 według odpowiedzi dotyczących poglądów, w tabeli nr 5 z liczebności całej grupy (N=78).

Tabela 3. Procenty według kategorii zmiennej niezależnej

procenty w wierszu dają razem 100,0%

| | zgadzam się | niezdecydowany | nie zgadzam się | razem |
|---|---|---|---|---|
| średnie | 12 | 6 | 28 | 46 |
| | 26,1 | 13,0 | 60,9 | 100,0 |
| licencjackie | 5 | 3 | 7 | 15 |
| | 33,3 | 20,0 | 46,7 | 100,0 |
| magisterskie | 11 | 2 | 4 | 17 |
| | 64,7 | 11,8 | 23,5 | 100,0 |
| razem | 28 | 11 | 39 | 78 |
| | 35,9 | 14,1 | 50,0 | 100,0 |

W tabeli nr 3 w wierszach obliczano procenty z sumy na końcu wiersza. W pierwszym wierszu liczebność 12 przedstawia 26,1% z sumy 46 (na prawym krańcu wiersza). Wszystkie procenty w tym wierszu dają razem 100,0% (26,1% + 13,0% + 60,9% = 100,0%). Tak samo oblicza się procenty we wszystkich wierszach.

Tabela 4. Procenty według kategorii zmiennej zależnej

procenty obliczone według odpowiedzi na pytanie ankietowe

procenty w kolumnie dają razem 100,0%

| | zgadzam się | niezdecydowany | nie zgadzam się | razem |
|---|---|---|---|---|
| średnie | 12 | 6 | 28 | 46 |
| | 42,9 | 54,5 | 71,8 | 59,0 |
| licencjackie | 5 | 3 | 7 | 15 |
| | 17,9 | 27,3 | 17,9 | 19,2 |
| magisterskie | 11 | 2 | 4 | 17 |
| | 39,3 | 18,2 | 10,3 | 21,8 |
| razem | 28 | 11 | 39 | 78 |
| | 100,0 | 100,0 | 100,0 | 100,0 |

W tabeli nr 4 w kolumnach obliczano procenty z sumy na końcu kolumny (na dole). W drugiej kolumnie liczebność 6 przedstawia 54,5% z sumy 11 (na dnie kolumny). Wszystkie procenty w tej kolumnie dają razem 100,0% (54,5% + 27,3% + 18,2% = 100,0%). Tak samo oblicza się procenty we wszystkich kolumnach.

Tabela 5. Procenty z liczebności całej grupy (N=78)

| | zgadzam się | niezdecydowany | nie zgadzam się | razem |
|---|---|---|---|---|
| średnie | 12 | 6 | 28 | 46 |
| | 15,4 | 7,7 | 35,9 | 59,0 |
| licencjackie | 5 | 3 | 7 | 15 |
| | 6,4 | 3,8 | 9,0 | 19,2 |
| magisterskie | 11 | 2 | 4 | 17 |
| | 14,1 | 2,6 | 5,1 | 21,8 |
| razem | 28 | 11 | 39 | 78 |
| | 35,9 | 14,1 | 50,0 | 100,0 |

procenty obliczone z liczebności całej grupy

procenty we wszystkich komórkach tworzą razem 100,0%

W tej tabeli w komórkach obliczano procenty z liczebności całej grupy (N=78). W pierwszej komórce (górna lewa) liczebność 12 stanowi 15,4% wszystkich respondentów N=78. Procenty we wszystkich komórkach tworzą razem 100,0% (15,4% + 7,7% + 35,9% + 6,4% + 3,8% + 9,0% + 14,1% + 2,6% + 5,1% = 100,0%). Procenty w kolumnie "razem" oraz w wierszu "razem" zostały też obliczone z N=78.

wpływ wykształcenia na poglądy

**Pierwsza tabela** (nr 3) pokazuje, jaki wpływ na poglądy ma wykształcenie. Z doświadczeń wiadomo, że ludzie z odmiennym wykształceniem mają odmienne poglądy, o które pytano w ankiecie. Ponieważ obie zmienne są powiązane, należy określić, która z nich jest zmienną niezależną, a która zależną. Można założyć, że w tej parze wykształcenie jest zmienną niezależną a poglądy zależną. Dlatego akurat pierwsza tabela jest najbardziej potrzebna. Przy badaniu związków (zależności) między zmiennymi praktycznie zawsze należy obliczać procenty w przedstawiony sposób.

wpływ poglądów na wykształcenie

**Druga tabela** (nr 4) pokazuje, jaki wpływ mają poglądy na wykształcenie. Z wyjątkiem bardzo nietypowych i rzadko spotykanych

przypadków taki kierunek wpływu jest bezsensowny. Dlatego też z takich tabel korzysta się rzadko.

włączanie niezwiązanych zmiennych do wspólnej tabeli

**Trzecia tabela** (nr 5) nie nadaje się do analizy związku między zmiennymi. Z tego powodu jest właśnie zupełnie niepotrzebna, ponieważ niepotrzebne jest włączenie niezwiązanych zmiennych do wspólnej tabeli.

W przypadku, gdy zmienne są powiązane, sensowna jest jedynie pierwsza tabela.

# Porządkowanie danych przedziałowych i ilorazowych

dane ilościowe porządkuje się na dwa sposoby

Uporządkowane dane nominalne lub porządkowe nadają się już do interpretacji. Inaczej jest w przypadku danych przedziałowych i ilorazowych. Te ostatnie dane porządkuje się w celach łatwiejszego opracowania. Korzystanie z komputera w opracowaniu danych przyniosło radykalną zmianę: porządkowanie danych ilościowych przed opracowaniem jest już zupełnie niepotrzebne. Dlatego też zagadnienie dotyczące porządkowania danych ilościowych omawiam w sposób skrócony. Dane ilościowe porządkuje się na dwa sposoby.

### 1. Tworzy się szereg szczegółowy prosty

szereg szczegółowy

Wartości zmiennej porządkuje się rosnąco. Sposób ten, bez szczegółowego opisu, zostanie zilustrowany jednym przykładem. Oto odpowiedzi nauczycieli na pytanie o staż pracy:

Tabela 6. Staż pracy nauczycieli

| x | 3 | 5 | 6 | 8 | 10 | 11 | 12 | 13 | 16 | 18 | 22 | 24 | 31 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |

### 2. Tworzy się szereg rozdzielczy z przedziałami klasowymi

szereg rozdzielczy

Tego sposobu używa się w sytuacjach, gdy liczebność grupy przewyższa liczebność zwykłej szkolnej klasy. Oto szereg rozdzielczy z wynikami uczniów z testu:

Tabela 7. Wyniki uczniów z testu

| przedziały klasowe | liczebność | procent liczebności | liczebność skumulowana | skumulowany procent liczebności |
|---|---|---|---|---|
| 8-12 | 2 | 1,7 | 2 | 1,7 |
| 13-17 | 5 | 4,2 | 7 | 5,8 |
| 18-22 | 8 | 6,7 | 15 | 12,5 |
| 23-27 | 11 | 9,2 | 26 | 21,7 |
| 28-32 | 17 | 14,2 | 43 | 35,8 |
| 33-37 | 21 | 17,5 | 64 | 53,3 |
| 38-42 | 15 | 12,5 | 79 | 65,8 |
| 43-47 | 15 | 12,5 | 94 | 78,3 |
| 48-52 | 12 | 10,0 | 106 | 88,3 |
| 53-57 | 10 | 8,3 | 116 | 96,7 |
| 58-62 | 4 | 3,3 | 120 | 100,0 |
| | 120 | | | |

szereg rozdzielczy z przedziałami klasowymi

# Przygotowanie danych do opracowania komputerowego

W empirycznych badaniach edukacji badacz dysponuje zwykle stosunkowo prostymi danymi. Najczęściej występujące zmienne to zmienne nominalne i porządkowe. Należą do nich m. in.: płeć, narodowość, wykształcenie, zadowolenie, powodzenie, różne zdolności, życzenia, zainteresowania itd. Rzadziej spotyka się zmienne przedziałowe lub ilorazowe. Zmienne przedziałowe lub ilorazowe w tym podrozdziale (w celu łatwiejszego zrozumienia) zostaną omówione wspólnie: wiek, wzrost i ciężar ciała, liczba uczniów w klasie, czas uczenia się, liczba przeczytanych książek itd. Niektóre zmienne, które należą do porządkowych, często opracowuje się tak, jak zmienne przedziałowe (np. wyniki uczniów w teście).

zmienne nominalne i porządkowe

zmienne przedziałowe i ilorazowe

jeden instrument badawczy dla wszystkich respondentów

Dla wszystkich jednostek dane zbiera się najczęściej za pomocą jednego instrumentu. Do najczęściej używanych instrumentów w badaniach pedagogicznych należą kwestionariusze ankiety, kwestionariusze wywiadu, skale postaw, skale szacowania, testy socjometryczne, testy wiadomości i protokoły obserwacji. W niniejszym opracowaniu osoby badane różnymi narzędziami określa się pojęciem „respondent".

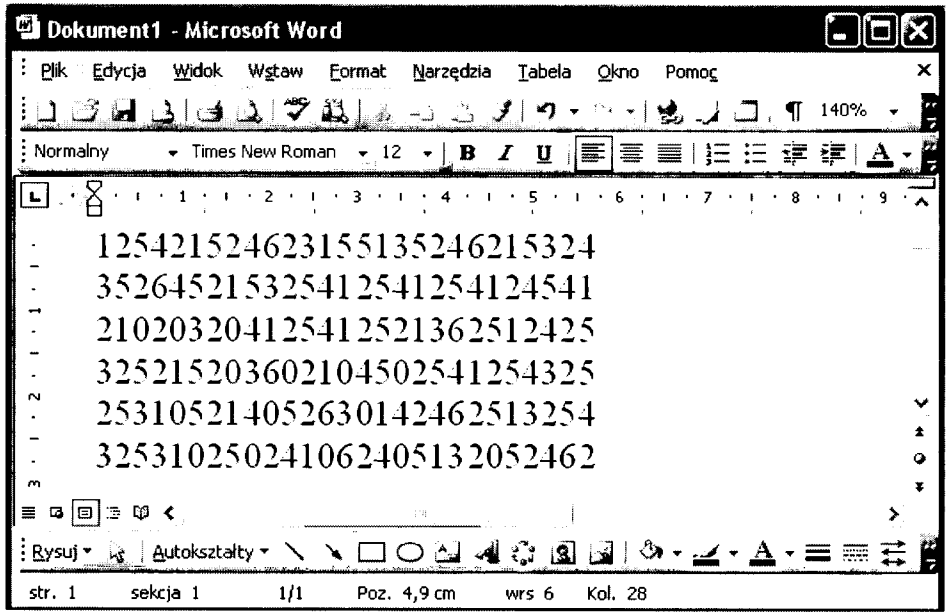
Na końcu badania dla wszystkich respondentów istnieje taka sama ilość danych (zmiennych), ponieważ dane zbiera się tymi samymi instrumentami dla wszystkich respondentów. To bardzo ułatwia wprowadzanie danych do komputera, ich opracowywanie i prezentowanie wyników.

dane każdego respondenta zapisuje się w nowym wierszu

Dane dla każdego respondenta zapisuje się w nowym wierszu. W ten sposób powstaje sieć danych.

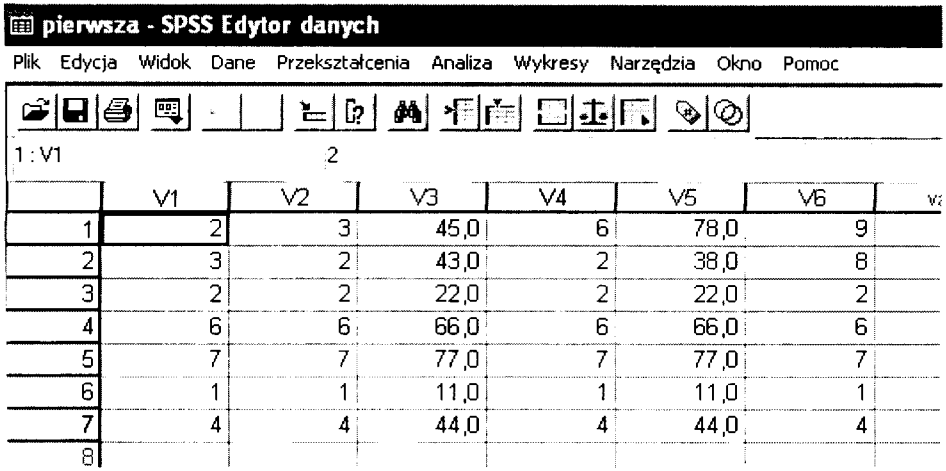
w wierszach umieszcza się respondentów

w kolumnach umieszcza się zmienne




Rysunek 1. Sieć danych w programie Word

W sieci danych w wierszach umieszcza się respondentów, natomiast w kolumnach zmienne. Podobna sieć powstaje także w programie



pierwsza - SPSS Edytor danych

Plik Edycja Widok Dane Przekształcenia Analiza Wykresy Narzędzia Okno Pomoc

1 : V1 2

| | V1 | V2 | V3 | V4 | V5 | V6 |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 45,0 | 6 | 78,0 | 9 |
| 2 | 3 | 2 | 43,0 | 2 | 38,0 | 8 |
| 3 | 2 | 2 | 22,0 | 2 | 22,0 | 2 |
| 4 | 6 | 6 | 66,0 | 6 | 66,0 | 6 |
| 5 | 7 | 7 | 77,0 | 7 | 77,0 | 7 |
| 6 | 1 | 1 | 11,0 | 1 | 11,0 | 1 |
| 7 | 4 | 4 | 44,0 | 4 | 44,0 | 4 |
| 8 | | | | | | |

sieć danych

Rysunek 2. Sieć danych w programie SPSS

program SPSS

Do opracowania takich danych używa się programów statystycznych, takich jak np. SPSS, Excel, itp.

## Spis zmiennych

pierwsza faza układania instrumentu

Spis zmiennych należy zrobić już przed układaniem instrumentu – jest to pierwsza faza układania instrumentu. Niektóre zmienne mają proste i „naturalne" imiona, takie jak: wiek, płeć, czas uczenia się, ilość błędów w dyktandzie, itd. Innym trudno nadać proste imiona i (przynajmniej na początku) nadaje się im "imiona techniczne" jak np. odpowiedź na czwarte pytanie w ankiecie, pierwsza ocena, druga ocena, pogląd pierwszy, itp. Program SPSS przyjmuje tylko imiona zmiennych nie dłuższe niż osiem znaków, dlatego dla wszystkich dłuższych imion należy przygotować odpowiedni skrót. Prawie zawsze ilość zmiennych dla wszystkich respondentów jest taka sama.

imiona zmiennych nie dłuższe niż osiem znaków

## Zapisy wartości

liczby jednocyfrowe

Wartości większej części zmiennych z ankiet, skal szacowania, protokołów obserwacji, itd. zostają zapisane w postaci liczb jednocyfrowych.

do komputera nie wnosi się liter

zawsze tylko liczby

Zilustruje to przykład pytania ankietowego zamkniętego z sześcioma proponowanymi odpowiedziami: A, B, C, D, E i F. Do komputera nie wnosi się liter – zawsze tylko liczby. Odpowiedź A zapisana zostanie jako liczba 1. Dla respondentów, którzy zakreślili odpowiedź B – zapisana zostanie liczba 2, dla wszystkich, którzy zakreślili odpowiedź C, zostanie zapisana liczba 3 i tak dalej do odpowiedzi F. Taka zmienna będzie miała jednocyfrowy zapis i będzie zajmowała w szeregu danych tylko jedno miejsce. W sieci danych będzie to jedna kolumna.

Oto przykład, w którym respondenci odpowiadali na pięć takich pytań.

pięć zmiennych

Pierwszy respondent zakreślił przy pytaniu pierwszym odpowiedź C, przy drugim A, przy trzecim A, przy czwartym D i przy piątym C. Odpowiedzi drugiego respondenta to kolejno: B, C, A, F, B; trzeciego: B, B, B, A, D i czwartego: C, C, E, A, F.

Należy to zapisać w następujący sposób:

pięć kolumn

| | |
|---|---|
| pierwszy respondent | 31143 |
| drugi respondent | 23162 |
| trzeci respondent | 22214 |
| czwarty respondent | 33516 |
| itd. | |

Dla pięciu zmiennych istnieje pięć kolumn, ponieważ wszystkie zapisy są jednocyfrowe.

Nie ma znaczenia, że wszystkie pytania w ankiecie nie są ilościowe (numeryczne). W ten sposób zapisuje się dane także dla wyłącznie jakościowych zmiennych. W przypadku płci należy np. dla kobiet zapisać liczbę 1, dla mężczyzn liczbę 2. Takie postępowanie nie zmienia natury zmiennych. Nie oznacza to, że nie uwzględnia się właściwości zmiennych. Łatwiejsze jest jedynie wprowadzanie danych do komputera, a także ich późniejsze opracowanie.

proste rozwiązania są najlepsze

Proste rozwiązania są najlepsze i przynoszą najmniej błędów. Dlatego też poszczególnym odpowiedziom na pytania w ankiecie (kategoriom zmiennej) przyporządkowuje się liczby w tej samej kolejności, według której były podane w ankiecie (pierwszej odpowiedzi zawsze 1, drugiej 2, trzeciej 3, itd.). Przy takim postępowaniu możliwość wystąpienia błędów jest najmniejsza.

Nawet w przypadku, jeżeli w pytaniu ankietowym nr 8 odpowiedzi zostały umieszczone w następujący sposób:

| | | |
|---|---|---|
| A. zawsze | B. czasami | C. nigdy |

a w pytaniu nr 13:

| | | |
|---|---|---|
| A. nigdy | B. czasami | C. zawsze |

odpowiedź A należy zapisać jako 1

w obydwu sytuacjach odpowiedź A należy zapisać jako 1 (bez względu na to, że to dwie zupełnie przeciwne odpowiedzi). Istnieje wprawdzie możliwość oznaczania jednakowych odpowiedzi zawsze taką samą liczbą (bez względu na fakt, czy były w ankiecie na pierwszym, drugim, trzecim... miejscu), jednak powoduje to komplikacje i zwiększa ilość błędów. Takie dylematy należy przewidzieć i rozwiązywać wcześniej – przy układaniu pytań ankietowych.

ilość błędów

W przypadku, gdy proponowanych odpowiedzi jest więcej niż dziewięć, nie można już zapisać odpowiedzi za pomocą jednej cyfry – konieczny jest zapis dwucyfrowy. Zakreślona pierwsza odpowiedź zostaje zapisana jako 01, druga odpowiedź jako 02, dziewiąta jako 09, dziesiąta jako 10, jedenasta jako 11, itd. Podobnie jest ze zmiennymi ilościowymi. Jeżeli w pytaniu o staż pracy pojawiają się odpowiedzi w ilości ponad dziewięć, należy dla wszystkich respondentów zapisać odpowiedzi w postaci dwucyfrowej liczby: 01, 02,.. 09, 10, 11,... 23, 24, itd. Są to zmienne, takie jak: rezultaty na testach, wiek, liczba błędów w dyktandzie, godziny ponadwymiarowe, itd. W przypadku, gdy najwyższy wynik testu wynosi np. 105 punktów, należy użyć dla wszystkich odpowiedzi zapisu trzycyfrowego: 002, 036, 067, 101, 105 itd. Dlatego też niektóre zmienne w sieci mają dwie, trzy lub jeszcze więcej kolumn.

zapis dwucyfrowy

Powstaje jednak pytanie: dlaczego nie można pierwszej odpowiedzi zapisać jako 1 a dwunastej jako 12? Odpowiedź brzmi: można, ale dużo prostsze jest tworzenie jednakowo szerokich kolumn przez dodawanie zer niż wprowadzanie danych w sposób pozornie prosty. Zapis z dodawaniem zer nazywa się zapisem o stałej szerokości. Takim sposobem zapisuje się dane po kolei bez odstępów (np. 52014206...). Przy zapisie bez dodawania zer należy rozgraniczać zmienne np. za pomocą przecinka (np. 5,2,1,4,2,6,...).

niepotrzebne komplikacje

Tabela 8. Dane przykładowego respondenta

zapisy liczbowe

| spis zmiennych | wartość lub odpowiedź | zapis liczbą | przedział możliwych wartości |
|---|---|---|---|
| ocena | bardzo dobra | 5 | 1-6+0 |
| płeć | mężczyzna | 2 | 1-2+0 |
| wyniki testu | 5 punktów | 5 | 0-145 |
| kierunek studiów | odpowiedź D | 4 | 1-7+0 |
| tryb studiów | zaoczne | 2 | 1-2+0 |
| staż pracy | 6 lat | 6 | 0-40 |

kolumny o stałej szerokości

dane oddzielone specjalnym znakiem

zapisy o stałej szerokości

Tabela 9. Zapis danych

| | |
|---|---|
| 520054206 | dane zawierają się w kolumnach o stałej szerokości |
| 5,2,5,4,2,6 | dane oddzielone są od siebie specjalnym znakiem |

Interpretacja:

- Przy ocenach pojawiają się wartości od 1 do 6 i dodatkowo 0 w przypadku braku odpowiedzi.
- W przypadku płci pojawią się tylko 1 i 2 oraz dodatkowo 0, jeżeli brak odpowiedzi.
- Przy wynikach testu (w punktach) pojawiają się wartości od 0 do 145. Nasz respondent osiągnął 5 punktów. Wiedząc, iż pojawiają się wartości do 145 przewidziano trzy kolumny – stąd zapis 005. Przy zapisie separowanym wystarczy 5 i przecinek.
- Przy kierunku ukończonych studiów pojawiają się odpowiedzi od A do G, czyli zapis od 1 do 7 i dodatkowo 0 w przypadku braku odpowiedzi.
- Przy trybie studiów pojawia się tylko 1 i 2 oraz dodatkowo 0 w przypadku braku odpowiedzi.
- Przy stażu pracy pojawiają się wartości od 0 do np. 40. Dlatego zapis o stałej szerokości, to 06, a zapis separowany – to 6 z przecinkiem.

Zapisując dane za pomocą przecinka należy za każdą zmienną postawić przecinek. Wymaga to dużego nakładu pracy. Można założyć, że typowe badanie zawiera 50 zmiennych. Jeżeli wśród zmiennych są np. trzy, które wymagają dwucyfrowego zapisu (druga, trzecia i czwarta), a pozostałe jednocyfrowego, to zapisy bedą wyglądać następująco:

typowe badanie

| zapis o stałej szerokości |
|---|
| 20504012354236254125362541251245212541254215241254122 |

| zapis przy pomocy przecinka |
|---|
| 2,5,4,1,2,3,5,4,2,3,6,2,5,4,1,2,5,3,6,2,5,4,1,2,5,1,2,4,5,2,1,2,5,4,1,<br>2,5,4,2,1,5,2,4,1,2,5,4,1,2,2 |

Dla zmiennych w badaniach edukacyjnych bardziej przydatny i ekonomiczny jest zapis o stałej szerokości.

ekonomiczność zapisu

Z zapisu oddzielonego przecinkami korzysta się głównie w innych dziedzinach – np. przy zapisywaniu nazwisk w ewidencjach komputerowych. Jako przykład można wskazać respondentkę Monikę Psiuk z Łodzi, którą należy wprowadzić do listy. Ponieważ mogą się pojawić dłuższe imiona niż Monika, dłuższe nazwiska niż Psiuk i dłuższe nazwy miejscowości niż Łódz, należy pozostawić dużo zer przed każdym zapisem. Dla imienia wybrano 15 miejsc, dla nazwiska 20, dla miejscowości też 15.

ewidencje komputerowe

| zapis o stałej szerokości |
|---|
| 000000000MONIKA00000000000000PSIUK00000000000ŁÓDŹ |
| zapis za pomocą przecinka |
| MONIKA,PSIUK, ŁÓDŹ |

Wprowadzanie takiej liczbi zer jest wymagające i niepotrzebne. Czasami, nawet i tak wielki zapas miejsc może okazać się nie wystarczający. Mieczysławę Kołowrotkiewicz z Warszawy jeszcze można w ten sposób zapisać, ale trudności pojawiają się, gdy respondentka po ślubie do swojega nazwiska dodała nazwisko Kotarbińska i przemeldowała się do miejscowości Nowe Miasto Lubawskie.

| |
|---|
| 000000000MONIKA00000000000000PSIUK00000000000ŁÓDŹ |
| 0000MIECZYSŁAWA00000KOŁOWROTKIEWICZ0000000WARSZAWA |
| 0000MIECZYSŁAWAKOŁOWROTKIEWICZ-KOTANOWEMIASTOLUBAW |

nie wystarczający zapas miejsc

brak miejsca

Zabrakło miejsca na pełne nazwisko (nie zmieściło się RBIŃSKA) i na pełną nazwę miejscowości (nie zmieściło się SKIE). Takie przypadki wymagałyby większej liczby zer, aby zawsze wystarczyło miejsca na całkowity zapis. Dlatego przy danych tego rodzaju dużo bardziej

zapis separowany

przydatny jest zapis separowany, ponieważ zmienna liczy się od przecinka do przecinka, bez względu na liczbę zawartych w zapisie znaków. Komputer w tym przypadku „wie", że imieniem jest wszystko do pierwszego przecinka, nazwiskiem wszystko od pierwszego do drugiego przecinka, a miejscowością wszystko od drugiego przecinka do końca. Oto przykładowy zapis separowany:

```
MIECZYSŁAWA, KOŁOWROTKIEWICZ-KOTARBIŃSKA, NOWE MIASTO LUBAWSKIE,
```

zapis o stałej szerokości

W empirycznych badaniach edukacji prawie bez wyjątków korzysta się z zapisu o stałej szerokości.

planowanie liczby kolumn

Przy planowaniu liczby kolumn w przypadku pytań zamkniętych na ogół nie ma żadnych komplikacji. Dla każdego pytania z ilości proponowanych odpowiedzi wiadomo, czy będzie wystarczał zapis jednocyfrowy, czy będzie potrzebny zapis dwucyfrowy. W przypadku pytań z numerycznymi odpowiedziami należy szybko przejrzeć wszystkie kwestionariusze, aby ustalić, jaka najwyższa liczba pojawia się w odpowiedziach. Bez tej orientacji może się zdarzyć następująca sytuacja: na przykład na wiek zaplanowano dwa miejsca, a w trakcie wprowadzania danych znajdzie się respondent w wieku 105 lat. Tej odpowiedzi nie można zapisać za pomocą dwóch miejsc. Sytuacja ta wymagałaby poprawienia wprowadzonych danych wszystkich respondentów (należałoby wszystkie zapisy dwucyfrowe zmienić na trzycyfrowe).

szereg danych

W taki sposób powstanie szereg danych dla każdego respondenta. Jeżeli brak jakiejś odpowiedzi (np. respondent nie zakreślił żadnej odpowiedzi lub odpowiedział, ale nie da się rozpoznać, co zakreślił itp.), należy w tym miejscu zapisać 0 (przy dwucyfrowym zapisie oczywiście dwa zera, itp.). Taki zapis powoduje, że wiersze dla wszystkich respondentów są jednakowej długości.

więcej niż jedna odpowiedź

Narzędzia do zbierania danych nie zawsze są tak proste, jak powyżej. Problemy pojawiają się najczęściej w tych pytaniach ankietowych, w których można zakreślić jednocześnie więcej niż jedną odpowiedź.

jedna odpowiedź

Niektórzy respondenci zakreślają jedną odpowiedź, niektórzy dwie lub więcej niż dwie, natomiast niektórzy nie zakreślają żadnej. Stąd też dla niektórych respondentów pojawia się więcej danych, dla innych natomiast tylko jedna wartość. Jak zatem zapisać odpowiedzi, aby zachować prostotę zapisu i równocześnie umieścić wszystkie dane? Najprostsze rozwiązanie jest następujące: z **jednego** pytania ankietowego zrobić **więcej** zmiennych – tyle, ile jest proponowanych odpowiedzi. Powstałe w ten sposób wszystkie zmienne będą jednocyfrowe. W przypadku, gdy respondent zakreśli pewną odpowiedź zapisuje się 1, jeżeli jej nie zakreśli zapisuje się 0. Ilustruje to przykład pytania ankietowego z sześcioma proponowanymi odpowiedziami dla ośmiu respondentów. W przykładzie istniała możliwość zakreślenia dowolnej liczby odpowiedzi.

więcej odpowiedzi

brak odpowiedzi

### *Przykład*

Tabela 10. Odpowiedzi ośmiu respondentów na pytanie ankietowe

| respondenci | | | | | | | |
|---|---|---|---|---|---|---|---|
| pierwszy | drugi | trzeci | czwarty | piąty | szósty | siódmy | ósmy |
| (A) | A | (A) | (A) | A | (A) | A | A |
| B | B | (B) | B | (B) | (B) | B | B |
| C | (C) | (C) | (C) | (C) | (C) | C | C |
| (D) | (D) | (D) | D | D | (D) | D | D |
| E | (E) | E | (E) | E | E | E | E |
| (F) | F | F | F | (F) | F | (F) | F |

odpowiedzi respondentów

Należy to zapisać następująco:

| | |
|---|---|
| pierwszy respondent | 100101 |
| drugi respondent | 001110 |
| trzeci respondent | 111100 |
| czwarty respondent | 101010 |
| piąty respondent | 011001 |
| szósty respondent | 111100 |
| siódmy respondent | 000001 |
| ósmy respondent | 000000 |

zapis odpowiedzi

wiersze jednakowej długości

Ósmy respondent nie zakreślił żadnej odpowiedzi, więc wpisano sześć razy zero. W ten sposób powstały wiersze (szeregi) danych jednakowej długości dla wszystkich respondentów.

Zjawisko to występuje bardzo często i z tego powodu zostanie dodatkowo wyjaśnione. W ankiecie pytano respondentów: Jakie są motywy podjęcia przez Panią/Pana studiów na kierunku pedagogika? Respondenci mogli zakreślić więcej niż jedną odpowiedź. Oto proponowane odpowiedzi:

więcej niż jedna odpowiedź

A będę miał możliwość awansu zawodowego
B lubię pracować z dziećmi
C lubię pracę grupową
D moi rodzice są nauczycielami
E będę miał długie wakacje

jedno pytanie

kilka odpowiedzi

Zadano tylko **jedno** pytanie, jednak każdy respondent mógł udzielić **kilku** odpowiedzi. W celu wprowadzania i opracowywania danych pytanie to w myślach należy przekształcić w kilka podobnych pytań. Odpowiedzi traktuje się w taki sposób, jak gdyby kolejne pytania brzmiały:

"Czy Pani/Pan zdecydowała się na studia pedagogiki z powodu możliwości awansu zawodowego?"

**TAK NIE**

"Czy Pani/Pan zdecydowała się na studia pedagogiki, bo lubi pracę z dziećmi?"

**TAK NIE**

„Czy Pani/Pan zdecydowała się na studia pedagogiki, bo lubi pracę grupową?"

**TAK NIE**

itd.

Jeżeli respondent zakreślił A, to tak jakby na pierwsze zasymulowane w myślach pytanie udzielił odpowiedź TAK. W przypadku, gdy go nie zakreślił, to jakby odpowiedział NIE. Tak należy postępować od pierwszej do ostatniej proponowanej odpowiedzi.

jedna zmienna

Ten rodzaj pytań można traktować jako jedną zmienną tylko w przypadkach, jeżeli opracowuje się go samodzielnie (czyli bez poszukiwania związków z innymi zmiennymi). Przy badaniu związków i zależności z innymi pytaniami lub zmiennymi ten rodzaj pytań należy traktować w powyżej opisany sposób (jako więcej zmiennych). Nie można analizować np. różnic według płci dla całego pytania, lecz dla każdej odpowiedzi oddzielnie (jedną po drugiej). Czynność ta jest skomplikowana, jednak nie istnieje inne rozwiązanie. Jeżeli współzależności nie są ważne, do kwestionariusza można włączyć podobne pytania. Jeżeli jednak współzależności są ważne dla badania, należy postępować w następujący sposób:

więcej zmiennych

- w kwestionariuszu ograniczyć możliwość wyboru odpowiedzi do jednej (np. który, z podanych motywów, był **najważniejszy** przy wyborze studiów pedagogiki?),

najważniejszy motyw

- do kwestionariusza włączyć dwa podobne do siebie pytania (jedno z nich dotyczy wszystkich motywów i istnieje możliwość wyboru kilku odpowiedzi, drugie dotyczy jednego – zwykle najsilniejszego motywu).

dwa podobne pytania

wszystkie motywy

Jeszcze bardziej złożone przypadki nie będą tu omówione, pomimo że mogą pojawić się w badaniach. Książka koncentruje się na najczęściej spotykanych przypadkach.

## Bezpośrednie i pośrednie wprowadzanie danych

proste kwestionariusze

skomplikowane kwestionariusze

W przypadku prostych kwestionariuszy dane wprowadzane są bezpośrednio z kwestionariuszy do komputera, bez wcześniejszego ich przygotowania. W przypadku skomplikowanych kwestionariuszy lepiej wprowadzać dane najpierw z kwestionariusza na specjalne kartki (szablony). Każdy respondent posiada wówczas swoją kartkę lub przynajmniej swój wiersz. Z wypełnionych kartek dane wprowadza się do komputera. O tym, czy dane wprowadza się bezpośrednio czy pośrednio, decyduje doświadczenie i praktyka w posługiwaniu się komputerem. Przy bezpośrednim wprowadzaniu jest mniej pracy, ale jednocześnie pojawia się więcej błędów.

# Wprowadzanie danych

*Z tego rozdziału nauczysz się:*

- *wprowadzać dane do komputera*
- *zachowywać wprowadzone dane*

## Word

edytory tekstu

Istnieje kilka edytorów tekstu. Ten, kto korzysta z jednego edytora, szybko może przejść do pracy z innym. Word jest na pewno najczęściej używanym edytorem tekstu. Oprócz edytorów istnieją i prostsze programy do edytowania (Wordpad, Notatnik itd.).

Word

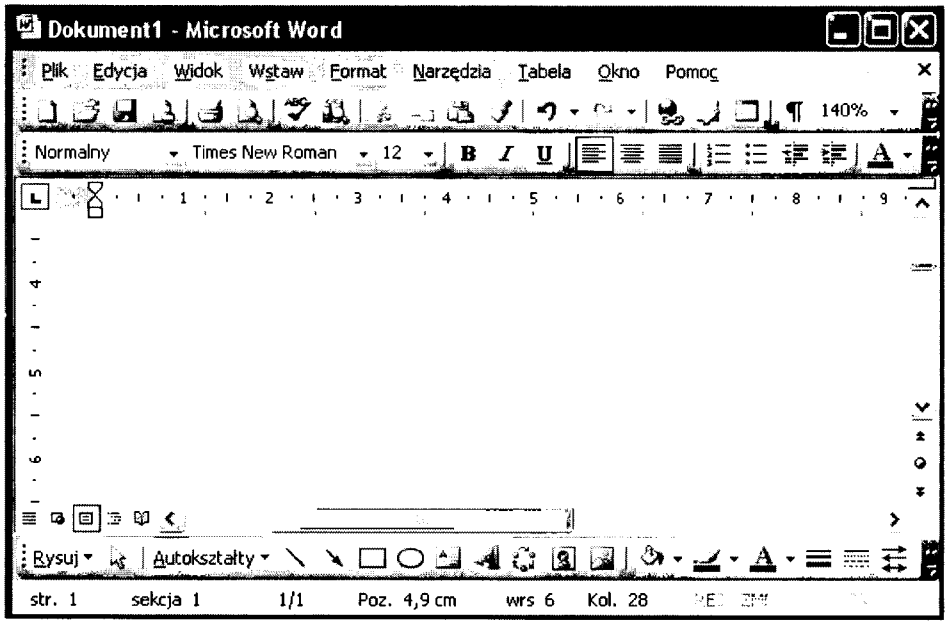
Czytelnik najprawdopodobniej już zna i korzysta z edytora Word. Dlatego też proponuję ten edytor do wprowadzania danych do komputera. Tak przygotowane dane można używać również w wielu innych programach.

## Okno programowe

uruchamianie programu Word

Kliknięciem na ikonę Word w systemie Windows uruchamia się program i otwiera okno programowe. Licząc, iż czytelnik korzysta już z programu Word, można od razu rozpocząć pracę w oknie.

okno programowe Word




Rysunek 3. Okno programu Word

pierwszy wiersz

Otwarty jest nowy dokument. W wiersze wnosi się dane pojedynczych respondentów. Każdy respondent ma swój wiersz. Dane zaczyna się wprowadzać kolejno, rozpoczynając od pierwszego respondenta. Do pierwszego wiersza wpisuje się dane (w postaci liczb) zwięzłe, czyli bez spacji. Ilość liczb jest nieograniczona, więc wiersz może być dowolnej długości. Jeżeli na ekranie zabraknie miejsca w pierwszym wierszu, Word sam przenosi się do wiersza drugiego, trzeciego itd. Na ekranie widać dane pierwszego respondenta w kilku wierszach, jednak dla programu SPSS jest to ciągle jeden wiersz (czyli pierwszy). Po wprowadzeniu danych pierwszego respondenta należy nacisnąć klawisz ENTER. Tym działaniem jednocześnie zaczyna się wypełnianie drugiego wiersza.

niepotrzebna komplikacja

Co prawda można postąpić inaczej: dane pierwszego respondenta rozdzielić (umieścić) w kilku wierszach korzystając z klawisza ENTER (np. po 50 cyfr w każdym wierszu). Jednak jest to niepotrzebna komplikacja. Zaawansowany użytkownik to wie, a dla innych jest prościej używać klawisza ENTER dopiero na końcu danych każdego respondenta.

wiersze utworzone przy użyciu ENTER

Ale uwaga: nieuważny użytkownik programu Word może nie rozpoznać, czy wiersze są tworzone przy użyciu klawisza ENTER, czy nie. Wiersze w obydwu przypadkach wyglądają jednakowo. Należy odróżniać „prawdziwe wiersze" (utworzone przy użyciu ENTER) i „pozorne wiersze" (stwarza je sam Word, gdy na ekranie braknie miejsca dla kolejnych cyfr). Jeżeli dane jednego respondenta były umieszczone w kilku wierszach za pomocą klawisza ENTER, każdy respondent będzie miał kilka wierszy.

kontrola, czy wszystkie dane zostały wniesione

Wracamy do wnoszenia danych w sposób prostszy (nie używa się klawisza ENTER). Każdy nowy respondent, to nowy wiersz. Długość wierszy jest jednocześnie kontrolą, czy wszystkie dane zostały wniesione. Jeżeli respondent nie odpowiedział na któreś z pytań, to w tym miejscu zapisuje się zero. Można używać też spacji, jednak pracy jest tyle samo, lecz spacje trudno policzyć. Najlepszym rozwiązaniem jest:

brak odpowiedzi

- bardziej doświadczony użytkownik powinien używać spacji; SPSS sam identyfikuje spacje jako brak danych (systemowy brak danych),

można używać spacji

- mniej doświadczony użytkownik powinien zapisywać zera: później należy programowi SPSS „powiedzieć", iż zera oznaczają brak danych – na szczęście jest to bardzo proste.

można używać zera

puste wiersze

Po zapisaniu danych ostatniego respondenta nie naciska się klawisza ENTER. Otwarty nowy wiersz oznaczałby nowego respondenta. Otwarty (i pusty) wiersz program SPSS traktuje jak respondenta, który nie udzielił żadnej odpowiedzi (bo wiersz jest pusty). Rzeczywista liczba respondentów w takim przypadku nie zgadza się z tym, co opracowuje SPSS. Należy na to szczególnie uważać. Jeżeli jednak enter został naciśnięty, należy za pomocą strzałki kursora iść na dół, aż do końca pliku. W momencie, gdy już nie można przesunąć niżej kursora,

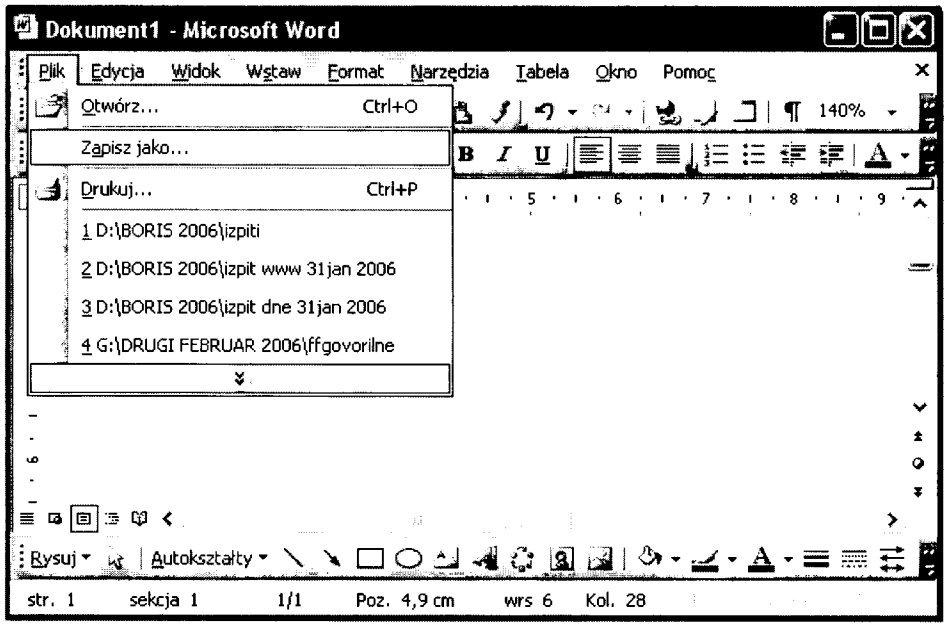
kasowanie pustych wierszy

należy skasować puste wiersze, korzystając z klawisza „strzałka w lewo" (Backspace, która znajduje się nad klawiszem ENTER). Z tym klawiszem trzeba presuwać kursor do ostatniego znaku końcowego wiersza. Posługiwanie się tym klawiszem wymaga jednak dużej uwagi. Jeżeli po naciśnięciu trzyma się go za długo, można skasować kilka znaków

krótkie przyciski

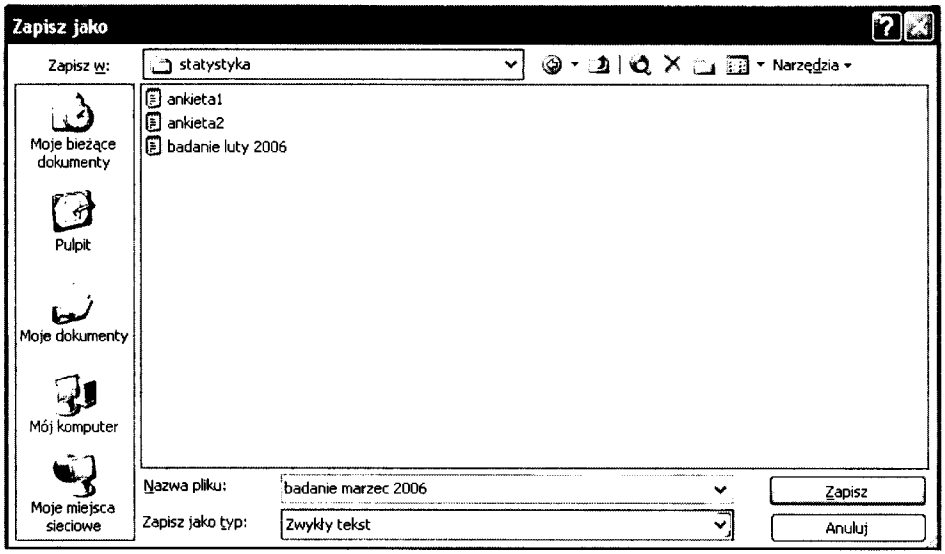
w sekundzie. Naciski muszą być krótkie (pojedyncze!!!). Na końcu tych poprawek należy zrobić sprawdzian: nacisnąć strzałkę kursora w dół. Jeżeli kursor nie przesuwa się niżej, oznacza to, iż niepotrzebne wiersze zostały skasowane. Wówczas można plik zapisać na dysk, czyli otworzyć menu Plik i wybrać pozycję Zapisz jako.

zapisywanie przez Zapisz jako




Rysunek 4. Polecenie Zapisz jako

Pojawi się okno Zapisz jako:




okno Zapisz jako

Rysunek 5. Okno Zapisz jako

Wybierz zapis w formacie **zwykły tekst** (lub plik tekstowy). Word automatycznie dodaje do nazwy pliku rozszerzenie TXT. Nic groźnego, jeżeli plik został błędnie zapisany w formacie Word-dokument (rozszerzenie DOC). Należy wówczas jeszcze raz otworzyć plik i ponownie zachować – teraz już w formacie TXT.

zwykły tekst

plik tekstowy

format TXT

Oznacza to koniec wnoszenia danych. Do opracowania danych w pliku można używać SPSS i innych programów. Dane są zapisane w sposób „uniwersalny". Uwaga: pliku DOC nie można opracowywać w SPSS!

format DOC

# Otwieranie i import danych

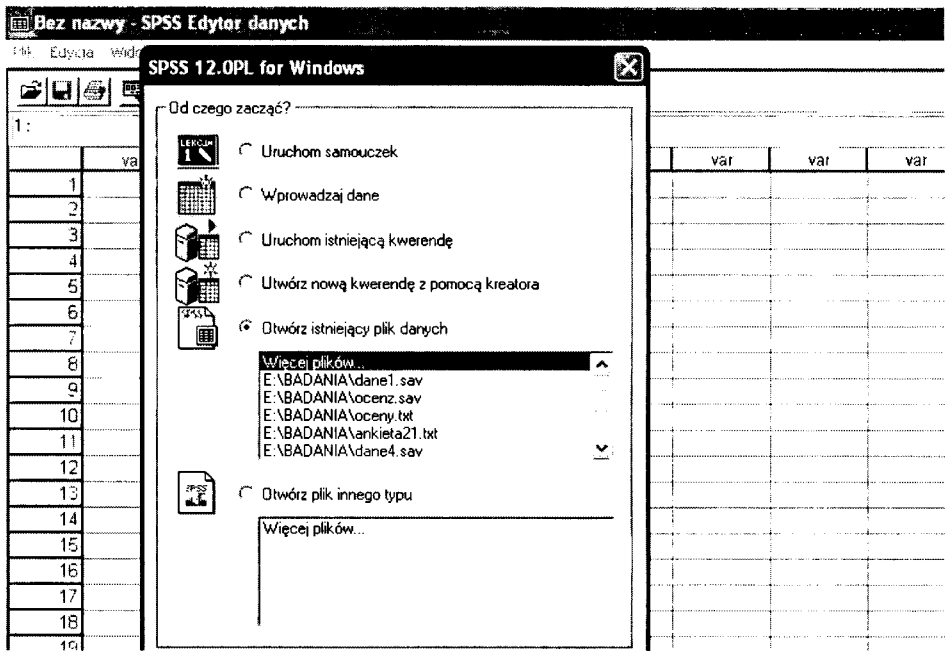
*Z tego rozdziału nauczysz się:*

- *importować dane z WORD do SPSS*
- *określać zmienne*
- *nadawać zmiennym i ich kategoriom dłuższe nazwy*
- *określać brakujące dane*

# Polecenia do otwierania i importu

Przy uruchamianiu programu SPSS otwiera się okno SPSS Edytor danych z dialogowym oknem początkowym:

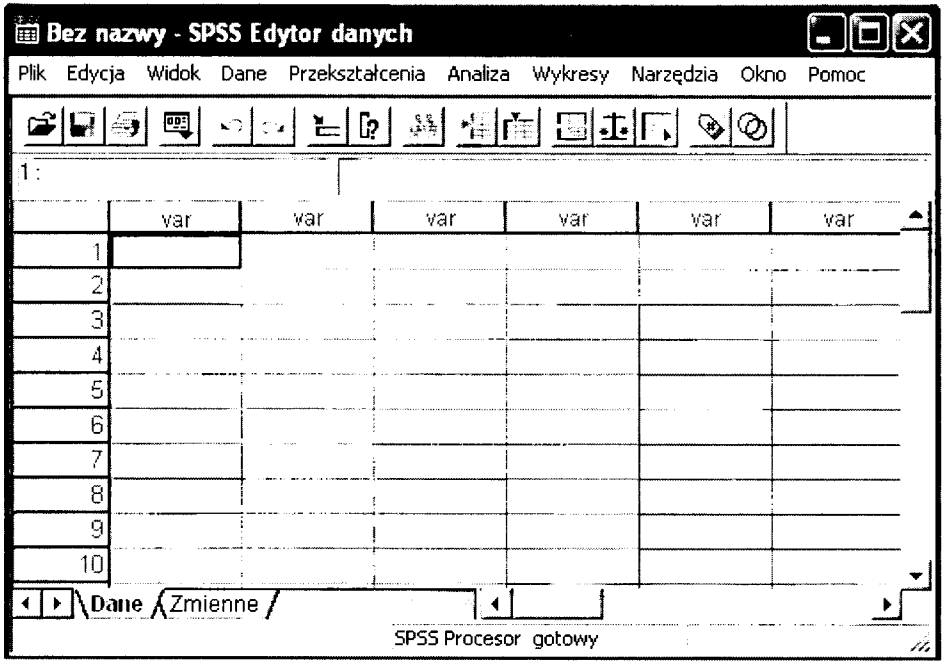
okno powitalne




Rysunek 6. Okno powitalne programu SPSS

importowanie danych

W tym oknie program SPSS proponuje kilka możliwości rozpoczęcia pracy. Ponieważ zamierzamy importować dane zapisane za pomocą programu Word, należy kliknąć na polecenie **Anuluj.** Na monitorze zostanie tylko okno Edytor danych.

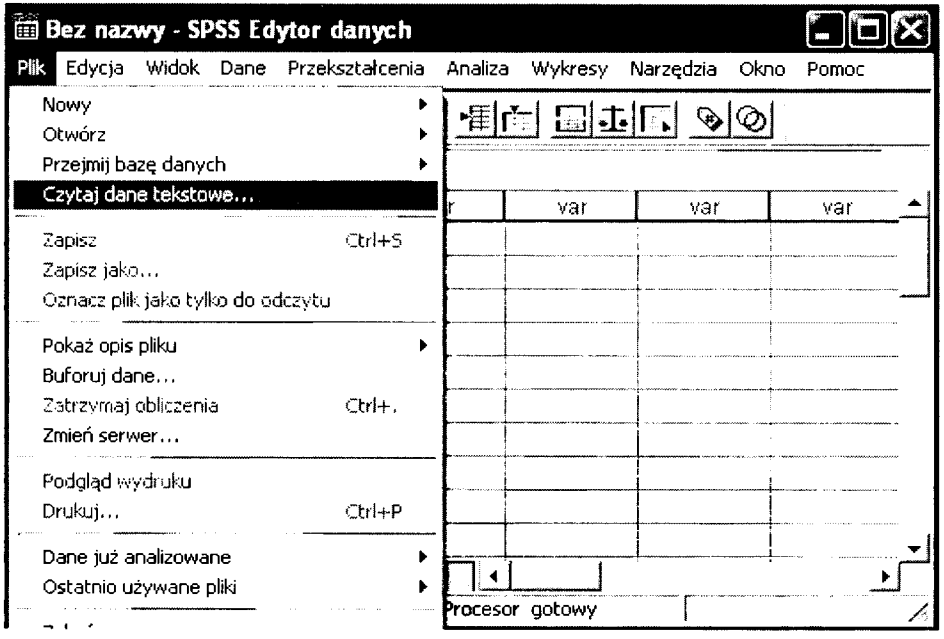



okno Edytor danych

Rysunek 7. Okno Edytor danych

Otwórz menu **Plik** i kliknij na pozycji **Czytaj dane tekstowe.**

otwieranie menu Plik

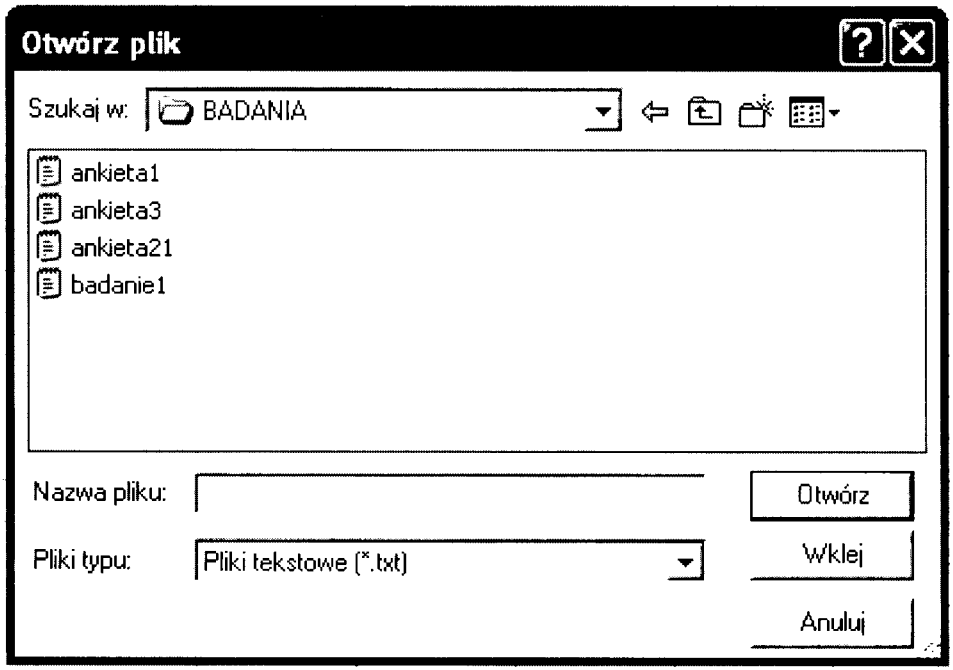



Czytaj dane tekstowe...

Rysunek 8. Polecenie Czytaj dane tekstowe

Otworzy się zwykłe okno do otwierania plików:

okno Otwórz plik



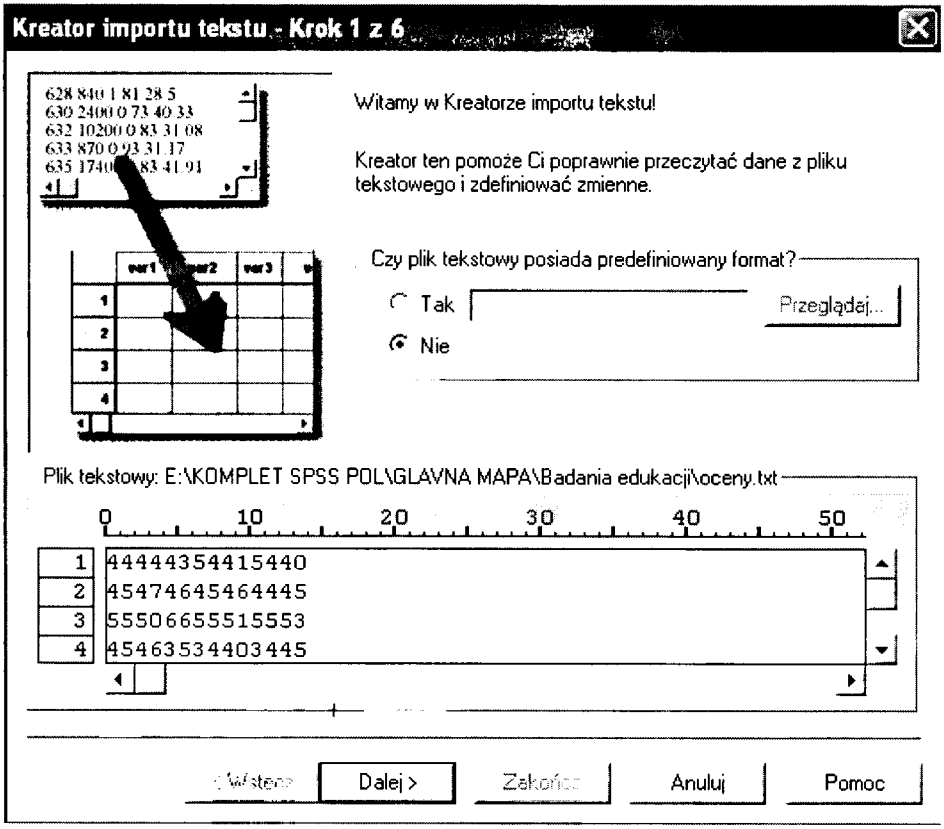

Rysunek 9. Okno Otwórz plik

pliki z rozszerzeniem TXT

uruchamianie Kreatora importu tekstu

Wybierz plik (z ikoną lub rozszerzeniem TXT) i kliknij na przycisk **Otwórz**. W ten sposób uruchamia się Kreator importu tekstu, który pomoże Ci poprawnie przeczytać dane z pliku tekstowego i zdefiniować zmienne. Program ten ma sześć kroków.

Oto okno kroku pierwszego:




okno Kreator importu tekstu krok pierwszy

czy dane są już zdefiniowane? odpowiedź Nie jest ustawiona automatycznie

Rysunek 10. Okno Kreator importu tekstu nr 1

Należy odpowiedzieć tylko na jedno pytanie: czy plik tekstowy posiada predefiniowany format? (czy dane są już zdefiniowane?) Dane przygotowane w programie Word nie są zdefiniowane, należy więc oznaczyć **Nie.** Odpowiedź ta jest już ustawiona automatycznie. Skontroluj i kliknij **Dalej.** Otworzy się okno kroku nr 2.

otwieranie okna kroku drugiego

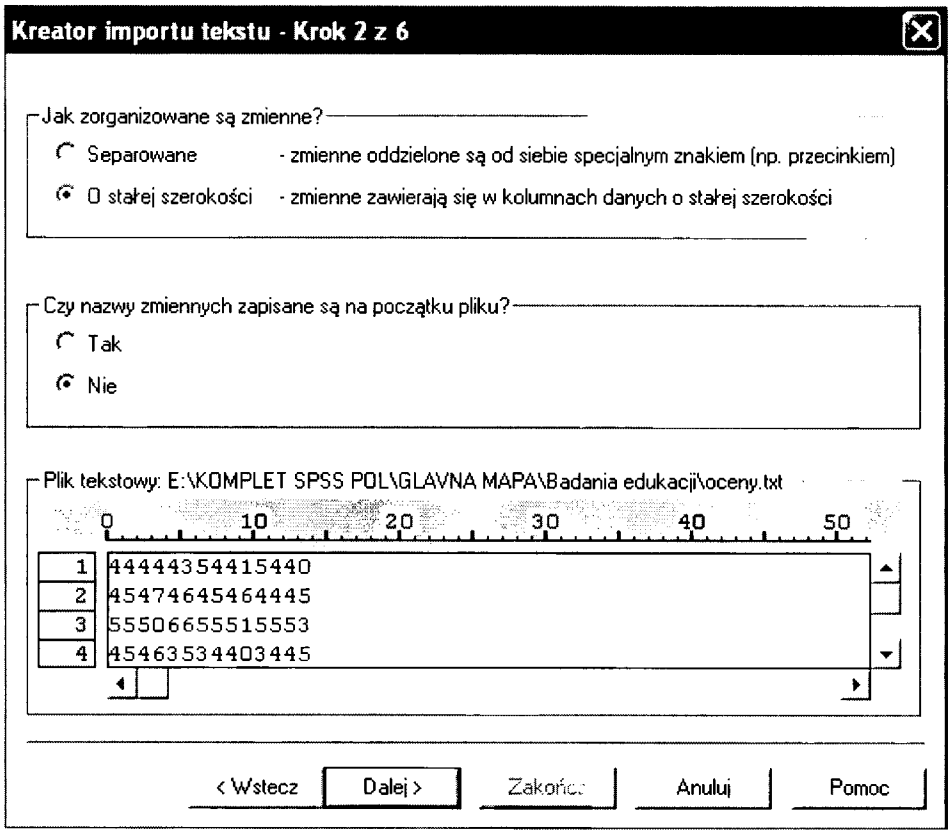
kliknij Dalej

Oto okno kroku drugiego:

okno Kreator importu tekstu krok drugi

odpowiedź „O stałej szerokości"

odpowiedź „Nie"



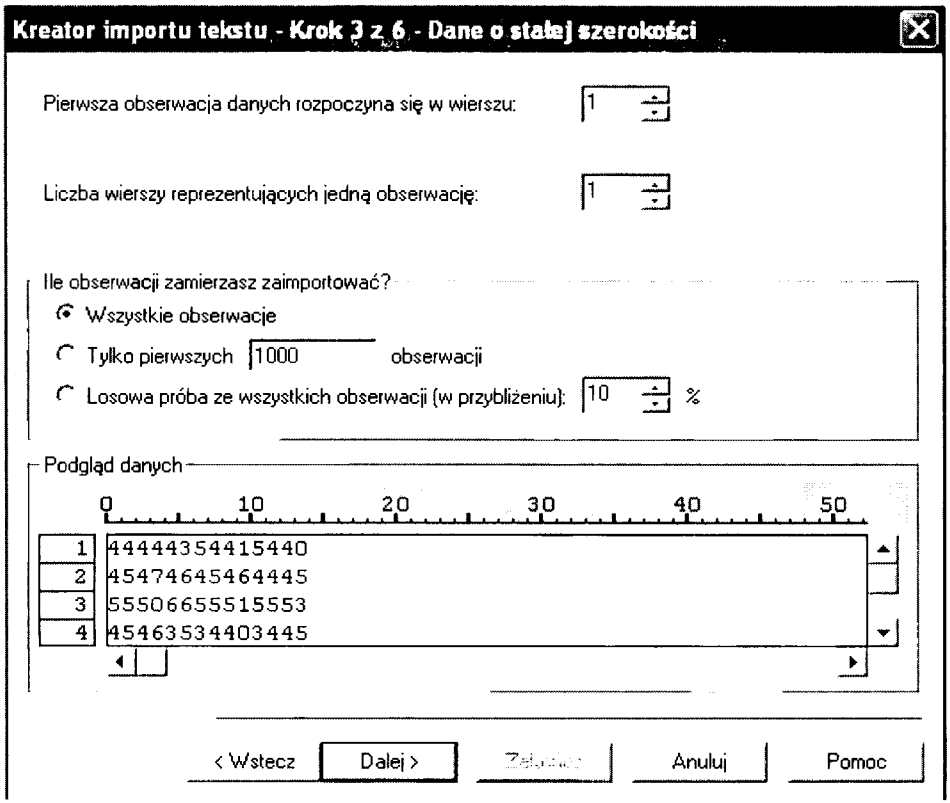

Rysunek 11. Okno Kreator importu tekstu nr 2

otwieranie okna kroku trzeciego

kliknij Dalej

Zostaw odpowiedzi **O stałej szerokości** i **Nie.** Odpowiedź „O stałej szerokości" oznacza, iż zmienne zawierają się w kolumnach o stałej szerokości (jednakowej szerokości dla wszystkich respondentów). Odpowiedź **Nie** odnosi się do pytania, czy nazwy zmiennych zapisane są na początku pliku. Tych nazw jeszcze nie ma. Kliknij **Dalej** i otworzy się okno Kreator importu tekstu kroku nr 3.

Oto okno kroku trzeciego:




okno Kreator importu tekstu krok trzeci

dwa razy odpowiedź „1"

odpowiedź „Wszystkie obserwacje"

Rysunek 12. Okno Kreator importu tekstu nr 3

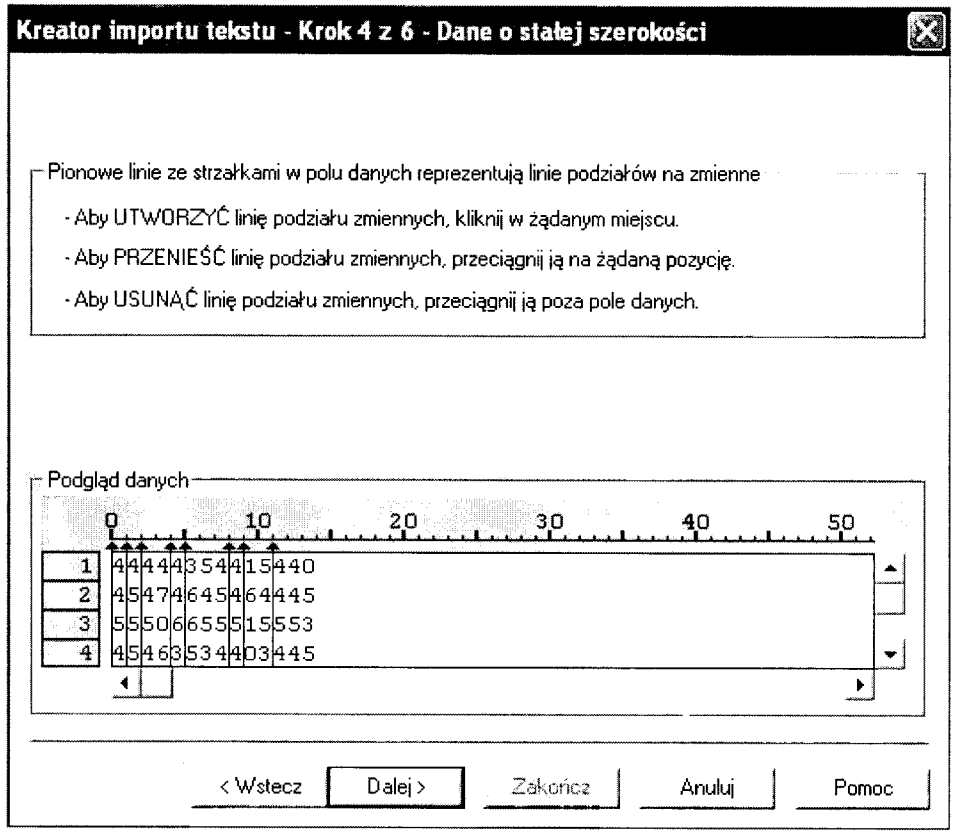
Wszystkie odpowiedzi są i tutaj ustawione w najprostszy sposób używania programu SPSS. Pierwsze pytanie brzmi: w którym wierszu rozpoczyna się pierwsza obserwacja danych? (czyli który wiersz zaczyna dane pierwszego respondenta?). W naszym przypadku pierwszy wiersz (czyli odpowiedź **1**). Kolejne pytanie brzmi: jaka jest liczba wierszy reprezentujących jedną obserwację? (czyli, ile wierszy ma każdy respondent?). W naszym przypadku tylko jeden wiersz, więc zostaw automatycznie ustawioną odpowiedź (**1**). W trzecim pytaniu SPSS chce dowiedzieć się, ilu respondentów ma uwzględnić. W naszym przypadku wszystkich, czyli należy wybrać pierwszą opcję **Wszystkie obserwacje.** W drugiej opcji można wybrać dowolną liczbę respondentów z początku listy. Trzecia opcja proponuje wybór losowej próby dowolnej wielkości (w %). Kliknij **Dalej** i otworzy się okno kroku nr 4.

otwieranie okna kroku trzeciego

kliknij Dalej

Oto okno do kroku czwartego:

okno Kreator importu tekstu krok czwarty




Rysunek 13. Okno Kreator importu tekstu nr 4

rozgraniczanie zmiennych

linie pionowe

To okno najważniejszego kroku i dlatego potrzebna jest odpowiednia uwaga. Tutaj rozgranicza się zmienne. Jeżeli pierwsza zmienna jest jednomiejscowa (jedna kolumna), to po pierwszej kolumnie należy postawić linie klikając myszką pomiędzy pierwszą i drugą kolumną. Stawianiem linii między odpowiednimi kolumnami określa się granice między zmiennymi. W przytoczonym przykładzie znajdują się kolejno zmienne: jednomiejscowa, jednomiejscowa, dwumiejscowa, jednomiejscowa, trzymiejscowa, jednomiejscowa, dwumiejscowa i trzymiejscowa. Linie można skasować naciskając myszką na mały trójkąt na górnym końcu linii. Na początku (przed pierwszą kolumną)

jest już automatycznie postawiona linia (właśnie niewidoczna – tego, że ona tam naprawdę jest, dowodzi trójkąt). Przy korekturach należy uważać, aby jej przypadkowo nie skasować! Od tej linii zaczynają się dane. W przypadkach, gdy na początku wiersza (każdego respondenta) jest kilka cyfr, które nie są danymi, linię należy usunąć. Są to rzadkie przypadki, w których przy wprowadzaniu danych na początku zapisano liczbę porządkową respondenta. Omawiany przypadek zilustrowano w tabeli nr 11.

pierwsza linia jest prawie niewidoczna

liczby porządkowe respondentów

Tabela 11. Dane z liczbami porządkowymi

| | |
|---|---|
| 0017252345..... | Dane zaczynają się dopiero od czwartej kolumny. Pierwsze trzy cyfry to liczba porządkowa (001 to pierwszy respondent, 002 drugi, 003 trzeci itd.). |
| 0024343253..... | |
| 0032352106..... | |
| 0045232323..... | |
| 00521541211..... | |
| 00622115214..... | |
| 00732541214..... | |
| 00851241212..... | |

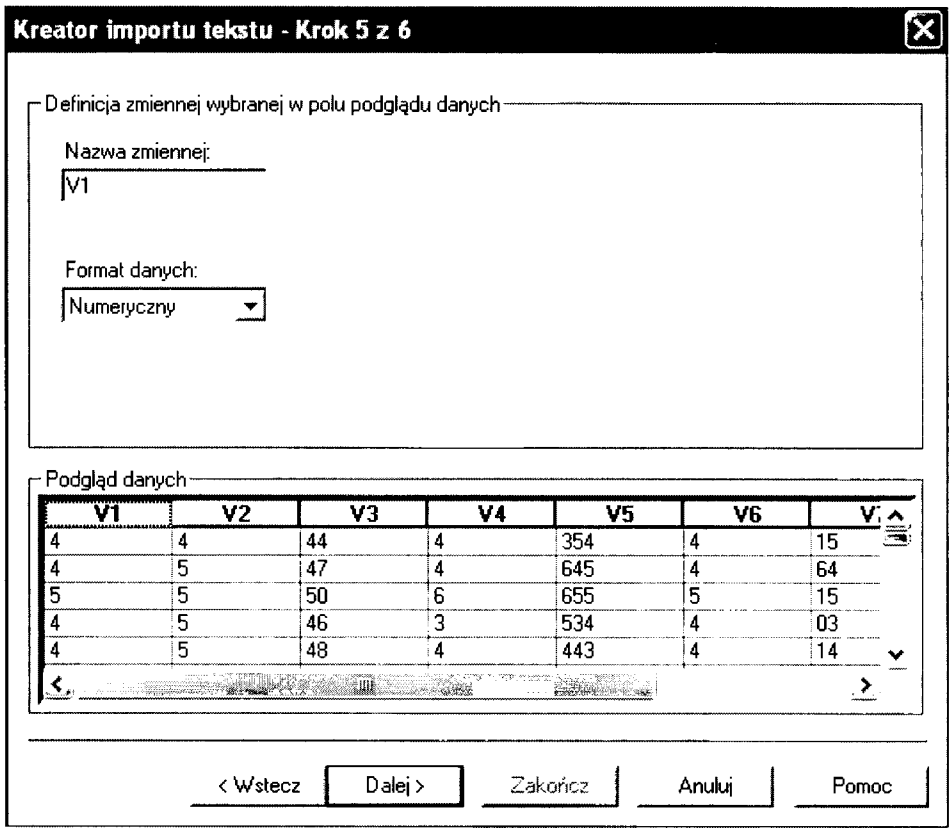
pierwsze trzy cyfry są niepotrzebne

W takim przypadku należy pierwszą pionową linię (nastawioną automatycznie) przestawić za trzecią cyfrą (pomiędzy trzecią i czwartą kolumną). Wtedy SPSS będzie pomijał pierwsze trzy cyfry, czyli zaczynał dopiero od czwartej kolumny. Przy wprowadzaniu danych niepotrzebne jest wprowadzanie liczb na początku. Nie będziemy zatem tego robić, ale wiedza ta jest nam jednak potrzebna, ponieważ przy opracowaniu danych innego badacza można się spotkać z takim przypadkiem. Na końcu ostatniej zmiennej nie wolno stawiać linii. Linia ta oznaczałaby, iż za ostatnią zmienną jest jeszcze jedna zmienna. Jeżeli wszystkie linie umieściłeś właściwie, kliknij **Dalej.** Otworzy się okno kroku piątego:

niepotrzebna linia na końcu

kliknij Dalej

okno Kreator importu tekstu krok piąty




Rysunek 14. Okno Kreator importu tekstu nr 5

nazwy zmiennych

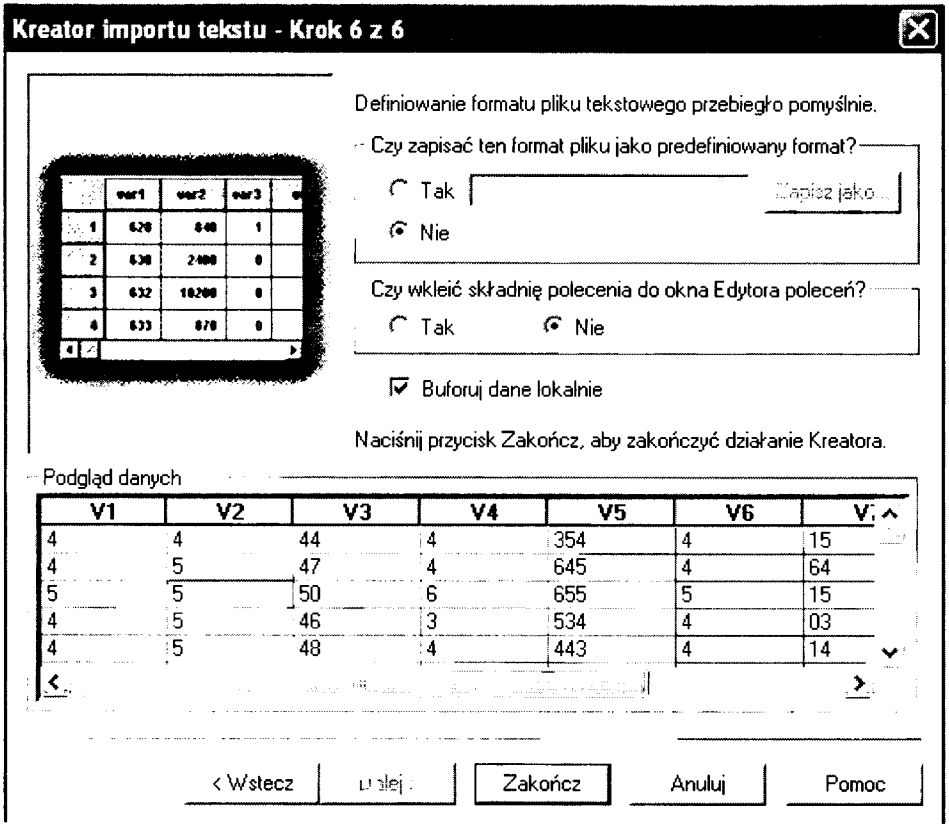
Tutaj dane są już podzielone na zmienne. Imiona zmiennych są „techniczne": V1, V2, V3 itd. (variable 1, variable 2...). Imiona te można zmienić i wstawić własne. Jednak nowe imiona mogą mieć maksymalnie osiem znaków (tylko litery i liczby, żadne inne znaki!). Nadawanie imion na tym etapie nie ma większego znaczenia. Lepiej później nadać zmiennym imiona pełne.

format danych

Kliknij na przycisku **V1** w podglądzie danych i wyświetli się pole „Nazwa zmiennej". Wpisz imię pierwszej zmiennej (np. wiek). Kliknij na V2, wpisz imię drugiej (np. wykształ, czyli skrót od imienia wykształcenie) i tak do ostatniej zmiennej. Można ustawić też rodzaj każdej zmiennej („Format danych"), ale to w naszym przypadku nie jest potrzebne.

Zakończ kliknięciem na przycisk **Dalej** i otworzy się okno kroku szóstego:

kliknij Dalej




okno Kreator importu tekstu krok szósty

odpowiedź „Nie"

odpowiedź „Nie"

pytania dla zaawansowanych

Rysunek 15. Okno Kreator importu tekstu nr 6

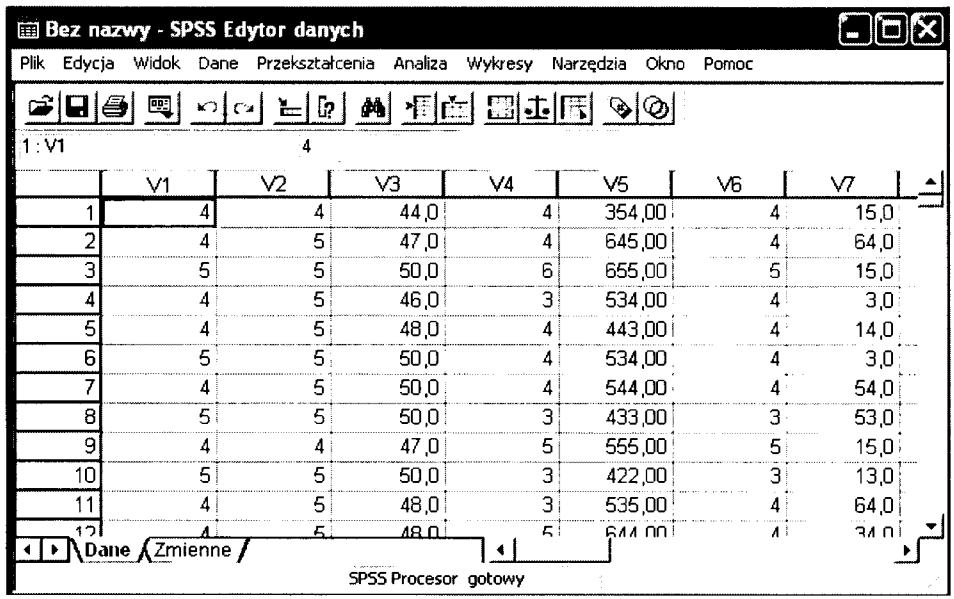
Pytania w kroku szóstym są już przewidziane dla zaawansowanych użytkowników SPSS, więc odpowiedz dwa razy **Nie.** Skontroluj, czy w poprzednich krokach wszystko dobrze zrobiłeś a jeżeli tak, kliknij na **Zakończ.**

kliknij na Zakończ

Po kilku sekundach pojawi się okno SPSS z siecią danych (należy być cierpliwym i nie klikać niepotrzebnie):

okno Edytor danych

widok na dane

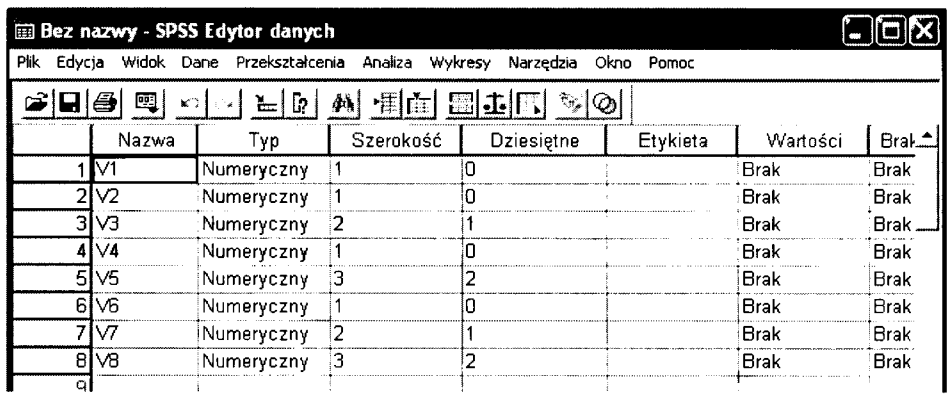



| | V1 | V2 | V3 | V4 | V5 | V6 | V7 |
|---|---|---|---|---|---|---|---|
| 1 | 4 | 4 | 44,0 | 4 | 354,00 | 4 | 15,0 |
| 2 | 4 | 5 | 47,0 | 4 | 645,00 | 4 | 64,0 |
| 3 | 5 | 5 | 50,0 | 6 | 655,00 | 5 | 15,0 |
| 4 | 4 | 5 | 46,0 | 3 | 534,00 | 4 | 3,0 |
| 5 | 4 | 5 | 48,0 | 4 | 443,00 | 4 | 14,0 |
| 6 | 5 | 5 | 50,0 | 4 | 534,00 | 4 | 3,0 |
| 7 | 4 | 5 | 50,0 | 4 | 544,00 | 4 | 54,0 |
| 8 | 5 | 5 | 50,0 | 3 | 433,00 | 3 | 53,0 |
| 9 | 4 | 4 | 47,0 | 5 | 555,00 | 5 | 15,0 |
| 10 | 5 | 5 | 50,0 | 3 | 422,00 | 3 | 13,0 |
| 11 | 4 | 5 | 48,0 | 3 | 535,00 | 4 | 64,0 |
| 12 | 4 | 5 | 48,0 | 5 | 644,00 | 4 | 34,0 |

Rysunek 16. Okno Edytor danych – widok na dane

W oknie tym można wybrać widok na sieć danych (Dane) lub widok na listę zmiennych (Zmienne). Oto widok na listę zmiennych:

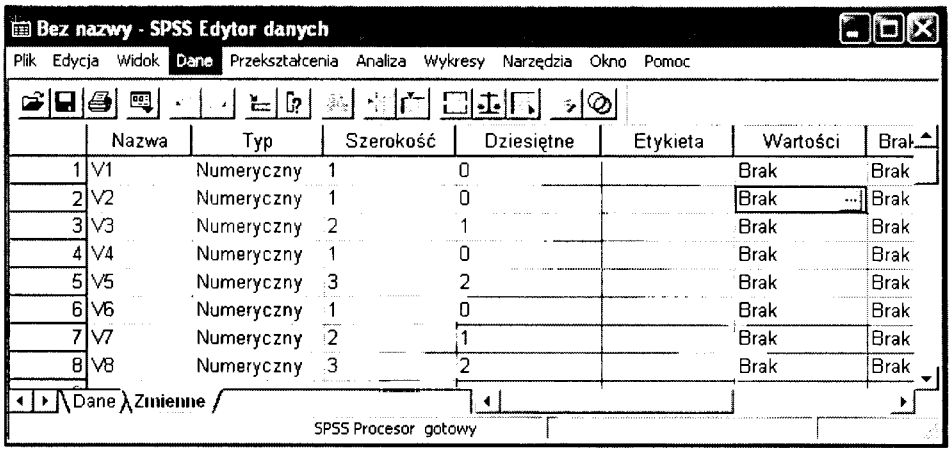
widok na zmienne




| | Nazwa | Typ | Szerokość | Dziesiętne | Etykieta | Wartości | Brak |
|---|---|---|---|---|---|---|---|
| 1 | V1 | Numeryczny | 1 | 0 | | Brak | Brak |
| 2 | V2 | Numeryczny | 1 | 0 | | Brak | Brak |
| 3 | V3 | Numeryczny | 2 | 1 | | Brak | Brak |
| 4 | V4 | Numeryczny | 1 | 0 | | Brak | Brak |
| 5 | V5 | Numeryczny | 3 | 2 | | Brak | Brak |
| 6 | V6 | Numeryczny | 1 | 0 | | Brak | Brak |
| 7 | V7 | Numeryczny | 2 | 1 | | Brak | Brak |
| 8 | V8 | Numeryczny | 3 | 2 | | Brak | Brak |
| 9 | | | | | | | |

Rysunek 17. Okno Edytor danych – widok na zmienne

Między oknami przechodzi się przyciskami Dane i Zmienne w lewym dolnym rogu. W oknie z siecią danych wybiera się poszczególne czynności opracowania statystycznego, a w oknie z listą zmiennych dodatkowo określa się zmienne.




widok na listę zmiennych

Rysunek 18. Okno Edytor danych – zmienne

# Dodatkowe określanie zmiennych

Szczególnie ważna jest lista zmiennych, gdyż tu można wszystko poprawić lub dopełnić. W niektórych przypadkach dodatkowe określenie zmiennych jest tylko korzystne, w innych natomiast koniecznie. Dla zmiennych charakterystycznych w badaniach edukacji najbardziej przydatne są trzy opcje: Etykieta, Wartości i Braki danych.

## *Etykieta*

dłuższe imiona

Opcja ta służy do nadawania zmiennym dłuższych imion. Pełne imiona są szczególnie ważne w momencie, gdy rezultaty przenosi się do programu Word w celu interpretacji. Raz nadane pełne imię zapisuje się wszędzie, gdzie pojawi się ta zmienna. Jako przykład może posłużyć zmienna, którą wprowadzono na początku pod imieniem np. ocenapol. W każdej tabeli z wynikami tej zmiennej będzie zapisane imię ocenapol. Po przeniesieniu tabeli do tekstu w programie Word, należy w każdej tabeli usunąć imię ocenapol i napisać pełne imię zmiennej np. końcowa ocena z języka polskiego w piątej klasie szkoły podstawowej. Dla doświadczonego użytkownika programu Word, być może nie przedstawia to większego problemu, ale i tak wymaga więcej pracy niż w przypadku, gdy korzysta się z opcji Etykieta nadając pełne imię zmiennej (raz na zawsze).

raz na zawsze

Opcje Etykieta uruchamia się w oknie z listą zmiennych:

kliknij na skrzyżowaniu kolumny Etykieta i wiersza zmiennej

wpisz pełne imię zmiennej

Bez nazwy - SPSS Edytor danych

Plik Edycja Widok Dane Przekształcenia Analiza Wykresy Narzędzia Okno Pomoc

| | Nazwa | Typ | Szerokość | Dziesiętne | Etykieta | Wartości | Braki danych |
|---|---|---|---|---|---|---|---|
| 1 | V1 | Numeryczny | 1 | 0 | | Brak | Brak |
| 2 | V2 | Numeryczny | 1 | 0 | | Brak | Brak |
| 3 | V3 | Numeryczny | 2 | 1 | | Brak | Brak |
| 4 | V4 | Numeryczny | 1 | 0 | | Brak | Brak |
| 5 | V5 | Numeryczny | 3 | 2 | | Brak | Brak |
| 6 | V6 | Numeryczny | 1 | 0 | | Brak | Brak |
| 7 | V7 | Numeryczny | 2 | 1 | | Brak | Brak |

Rysunek 19. Okno Edytor danych

Kliknij w okienku na skrzyżowaniu kolumny Etykieta i wiersza tej zmiennej, której chcesz nadać pełne imię. Od razu można wpisać pełne imię zmiennej. Następnie należy kliknąć w kolejnym okienku, wprowadzić imię dalszej zmiennej, itd. Oto przykład:

ocenz - SPSS Edytor danych

Plik Edycja Widok Dane Przekształcenia Analiza Wykresy Narzędzia Okno Pomoc

| | Nazwa | Typ | Szerokość | Dziesiętne | Etykieta | W |
|---|---|---|---|---|---|---|
| 1 | V1 | Numeryczny | 1 | 0 | udział ucznia w kol | Brak |
| 2 | V2 | Numeryczny | 1 | 0 | ocena z matematyk | Brak |
| 3 | V3 | Numeryczny | 1 | 0 | test AKNM-1 | Brak |
| 4 | V4 | Numeryczny | 1 | 0 | test AKNM-2 | Brak |
| 5 | V5 | Numeryczny | 1 | 0 | odpowiedz na pierw | Brak |
| 6 | V6 | Numeryczny | 1 | 0 | | Brak |
| 7 | V7 | Numeryczny | 1 | 0 | | Brak |
| 8 | V8 | Numeryczny | 1 | 0 | | Brak |
| 9 | V9 | Numeryczny | 1 | 0 | | Brak |

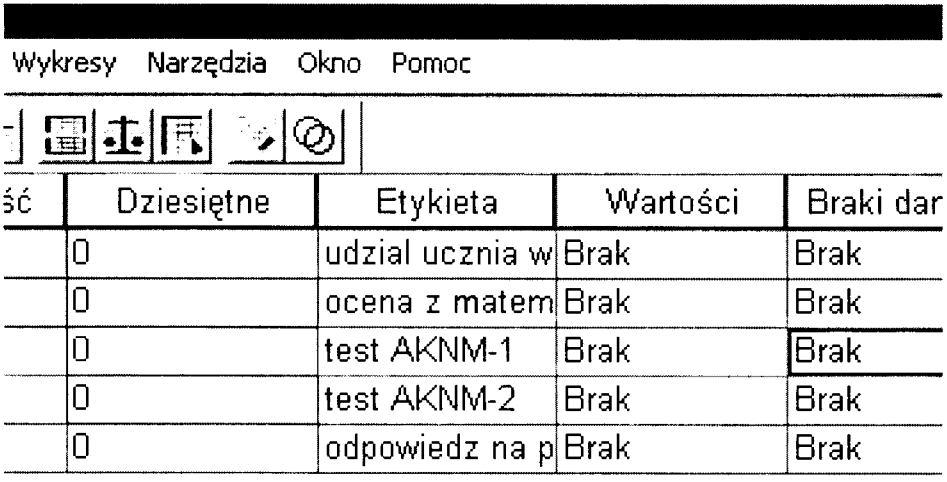
Rysunek 20. Okno Edytor danych – etykieta

można używać też polskich znaków

Imię nie powinno przekraczać 256 znaków. Można używać też polskich znaków: ą, ę, ł, ó, ź, itd. Te znaki, dopóki pracuje się w SPSS nie będą zapisywane poprawnie, ale po przeniesieniu tabel i wyników do tekstu w programie Word, będą zapisane poprawnie.

Okienko z imieniem zmiennej rozszerza całą kolumnę i przesuwa wszystkie kolumny po prawej stronie jeszcze bardziej w prawo. Po wprowadzeniu imienia można myszką chwycić prawą krawędź kolumny i kolumnę zwęzić. Na ekranie będzie widocznych tylko kilka pierwszych liter imienia, jednak imię już zostało wprowadzone.

można zwęzić kolumnę




| ść | Dziesiętne | Etykieta | Wartości | Braki dar |
|---|---|---|---|---|
| | 0 | udzial ucznia w | Brak | Brak |
| | 0 | ocena z matem | Brak | Brak |
| | 0 | test AKNM-1 | Brak | Brak |
| | 0 | test AKNM-2 | Brak | Brak |
| | 0 | odpowiedz na p | Brak | Brak |

Rysunek 21. Okno Edytor danych

## *Wartości*

Za pomocą tej funkcji można nadać imiona poszczególnym kategoriom zmiennej. W pliku danych wszystkie wartości są zapisane w postaci liczb. W przypadku zmiennych ilościowych zapis ten odpowiada rzeczywistości. Np. na pytanie, ile godzin uczeń się uczył, poprawną odpowiedzią jest faktycznie 5 lub 37 lub inna liczba. W przypadku zmiennych jakościowych, np. w pytaniu o płeć uczniów, liczby tylko zastępują prawdziwe odpowiedzi. W tabelach oczywiście muszą znajdować się prawdziwe kategorie, a nie zastępcze liczby. W pytaniu o płeć studentów są to kategorie kobieta i mężczyzna, a nie 1 i 2, w pytaniach ankietowych mają być umieszczone słowne odpowiedzi zamiast 1, 2, 3, 4, itd.

zmienne ilościowe

zmienne jakościowe

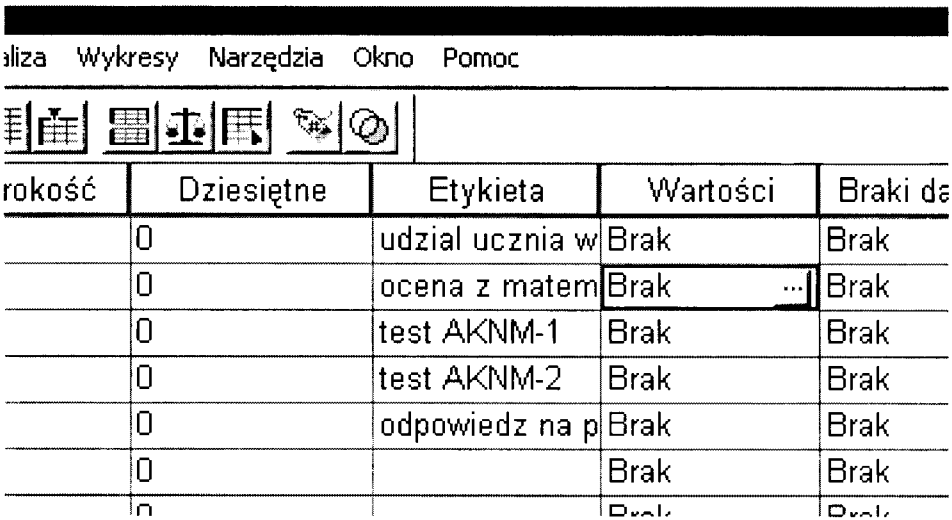
słowne odpowiedzi zamiast liczb

wartości
imiona kategorii

Kategoriom można nadać imiona używając funkcji Wartości. Kliknij w okienku na skrzyżowaniu kolumny Wartości i wiersza tej zmiennej, której chcesz nadać imiona kategorii. W okienku pojawi się szary kwadracik.

kliknij na skrzyżowaniu kolumny Wartości i wiersza zmiennej

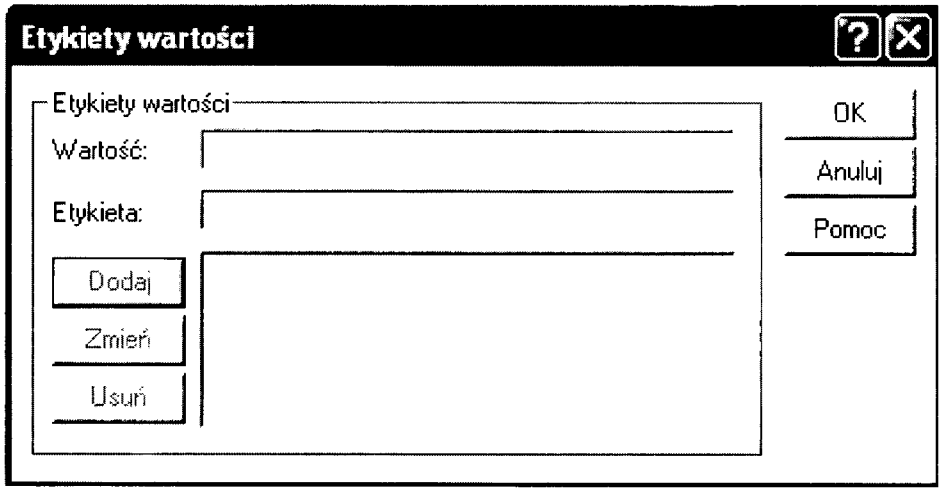
kliknij na kwadraciku




Rysunek 22. Okno Edytor danych

Kiknij na kwadraciku i otworzy się okno:

okno Etykiety wartości




Rysunek 23. Okno Wartości

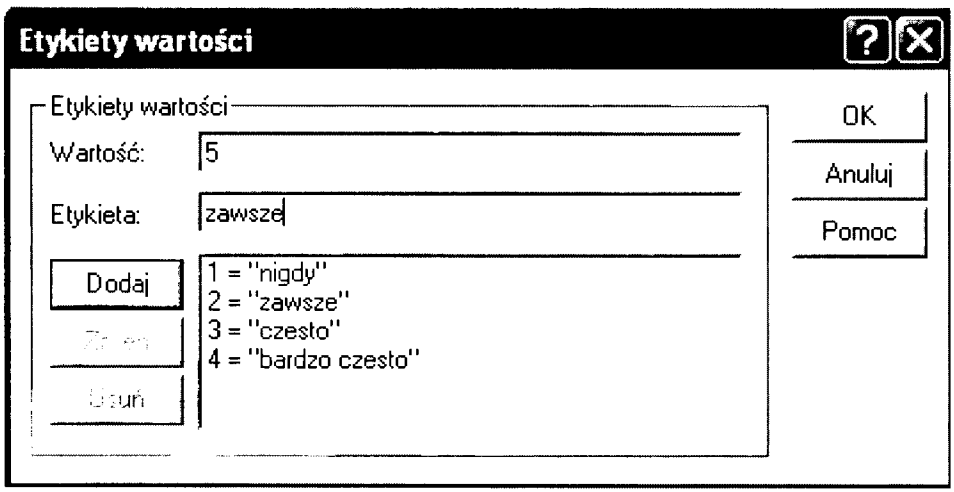
W polu Wartość kliknij i wpisz pierwszą z wartości (np. 1), a w Etykieta imię tej kategorii (np. kobieta). Następnie kliknij na przycisk **Dodaj**. Tym przyciskiem wpis potwierdza się i przenosi do listy nowych imion. Jeszcze raz kliknij w polu Wartość i wpisz drugą z wartości (np. 2), a w Etykieta imię drugiej kategorii (np. mężczyzna). Następnie kliknij na przycisk **Dodaj.** Po wprowadzeniu wszystkich wpisów kliknij na **OK.** Oto przykład przed kliknięciem na Dodaj i OK:

okno Etykiety wartości

wpisz pierwszą z wartości

wpisz imię tej kategorii

kliknij na przycisk Dodaj




Rysunek 24. Okno Etykiety wartości

Błędne wpisy można poprawić w każdej chwili. Na liście końcowych zmian kliknij na błędnym wpisie, a następnie w górnych rubrykach popraw wpis i potwierdź poprawkę przyciskiem **Zmień.** Wpis można też usunąć tak: kliknij na błędnym wpisie i następnie na **Usuń.**

kliknij na błędnym wpisie i następnie na Usuń

## *Braki danych*

Funkcja Braki danych umożliwia określenie brakujących danych. Przy wprowadzaniu danych w przypadku braku odpowiedzi wprowadza się zero. Program SPSS tego »nie wie« i należy mu to »powiedzieć«.

1. zero stanowi prawdziwą odpowiedź

2. zero oznacza brak odpowiedzi

***Uwaga!***

Na niektóre pytania zero stanowi prawdziwą (poprawną) odpowiedź! Na pytanie o płeć zero oznacza brak odpowiedzi, ale na pytanie, ile książek uczeń przeczytał, zero jest ważną odpowiedzią (prawdziwą, poprawną). W niektórych przypadkach badacz chce mieć w tabelach brakujące odpowiedzi w odrębnej kolumnie. W innych przypadkach te odpowiedzi należy zupełnie usunąć – i z opracowania i z tabeli. Właśnie do tego służy funkcja Braki danych. Wszystko, co należy zrobić to określić, dla których zmiennych zero jest brakiem odpowiedzi.

Wskazówki dla bardziej zaawansowanego użytkownika SPSS:

wskazówki dla bardziej zaawansowanych

Dla każdej zmiennej należy odrębnie określić, jak są oznaczone brakujące dane. Dla niektórych zmiennych jest to zero, dla drugich inny znak, dla innych wcale nie jest potrzebna ta funkcja. Niekiedy istnieje potrzeba odróżnienia braku odpowiedzi od niepoprawnych odpowiedzi. Niepoprawna odpowiedź to przypadek, w którym respondent udzielił odpowiedzi, ale nie ma możliwości jej jednoznacznego określenia (np. jeżeli zakreślił kilka odpowiedzi na pytanie, gdzie wymagana jest tylko jedna odpowiedź lub jeżeli odpowiedzi są zakreślone, następnie przekreślone i w końcu nie wiadomo której odpowiedzi respondent udzielił itd.). W tych przypadkach dla brakujących odpowiedzi wybiera się zero, a dla niepoprawnych jakiś inny znak. Wybiera się oczywiście znak, który nie pojawia się wśród odpowiedzi. Zwykle jest to liczba 9, ponieważ wyjątkowo rzadko spotyka się pytania ankietowe z dziewięcioma kategoriami. Wówczas dla obydwu znaków (0 i 9) używa się Braki danych, a w tabelach otrzymuje się informacje o tym, ile jest jednych, a ile drugich odpowiedzi (oczywiście obok wszystkich kategorii ważnych odpowiedzi).

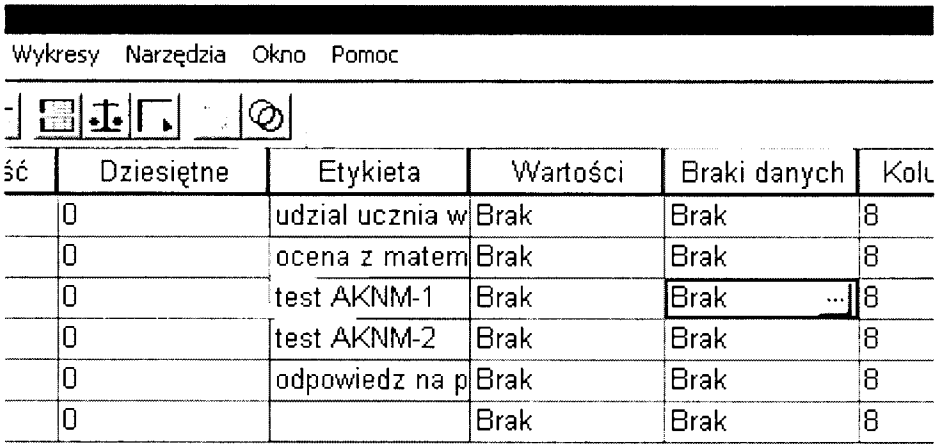
SPSS traktuje spacje jako brak danych

Należy nadmienić, iż program SPSS automatycznie traktuje spacje jako brak danych (spacja jest odbierana jako systemowy brak danych). Jeżeli użytkownik w miejscu brakujących danych wpisuje spacje, wówczas nie musi wcale korzystać z funkcji Braki danych (wszystko odbędzie się automatycznie). Jednak korzystanie ze spacji ma dwie wady:

- Spacje są trudne do policzenia. Przy wprowadzaniu danych, w przypadkach, gdzie należy wpisać kilka spacji po kolei, trudno rozpoznać, ile już ich wpisano, a ile jeszcze należy wpisać.
- Przy importowaniu danych należy być bardzo uważnym. Jeżeli w bazie danych znajduje się choć jedna spacja, SPSS odbiera ją jako znak rozgraniczania zmiennych. Dane mogą być sporządzone bez spacji pomiędzy zmiennymi (O stałej szerokości) lub ze spacjami (Separowane). Po znalezieniu przynajmniej jednej spacji, w drugim kroku importu na pytanie o formacie, SPSS automatycznie ustawia odpowiedź Separowane. Jednak nie jest to poprawna odpowiedź, ponieważ naszym formatem jest format o nazwie O stałej szerokości. Po wybraniu odpowiedzi O stałej szerokości SPSS odbiera spacje jako brak danych. Nie jest to wprawdzie wielka komplikacja, ale i tak lepiej korzystać z zera zamiast ze spacji.

dwie słabości korzystania ze spacji

Aby określić brak danych, należy kliknąć w okienku na skrzyżowaniu kolumny Braki danych i wiersza tej zmiennej, której określa się brakujące dane. W okienku pojawi się szary kwadracik.




okno Zmienne

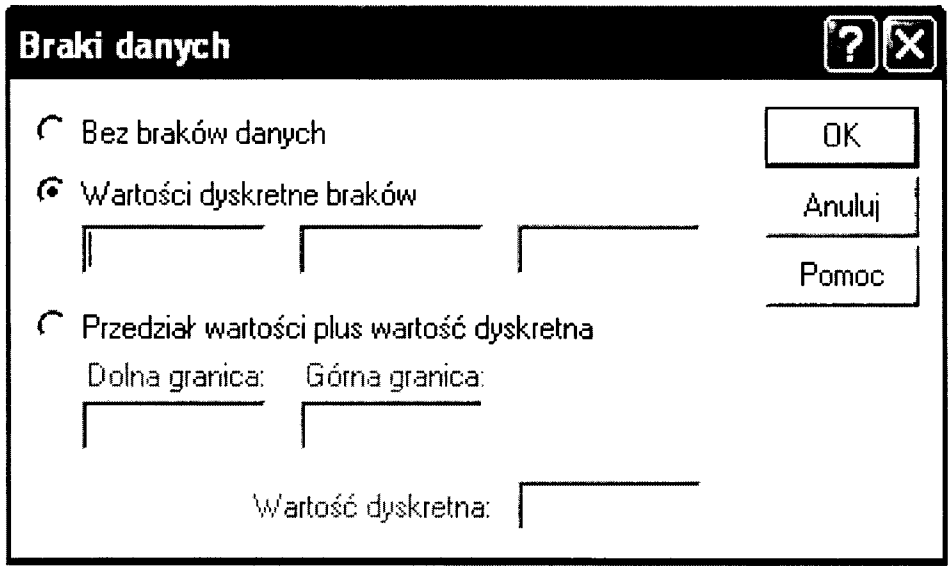
kliknij na skrzyżowaniu kolumny Braki danych i wiersza zmiennej

Rysunek 25. Okno Edytor danych

Kliknij na kwadracik i otworzy się okno Braki danych:

kliknij na kwadracik

okno Braki danych




Rysunek 26. Okno Braki danych

wartości dyskretne braków

- Jeżeli brakujące dane były wprowadzane jako jeden lub kilka ściśle określonych znaków, należy wybrać rubrykę Wartości dyskretne braków i w puste pole wpisać znak (jeżeli znaków jest kilka, to w każdym polu umieszcza się jeden znak). Zwykle jest to tylko zero, a pozostałe pola zostają puste. Jest to wystarczające prawie we wszystkich przypadkach.

przedział wartości plus wartość dyskretna

- Jeżeli brakujące dane określa się jako wartości »od...do«, wtedy wybiera się rubrykę Przedział wartości plus wartość dyskretna oraz w lewe pole wpisuje się dolną granicę, a w prawe pole górną granicę (a w dodatkowym oknie wartość dyskretną, jeżeli taka istnieje).

kliknij na OK

Po wprowadzeniu wartości należy kliknąć na **OK.**

Opisane sposoby dodatkowego porządkowania i określania danych w pełni wystarczą w badaniach edukacji. Możliwości programu SPSS są

dużo większe, ale książka, która opisywałaby wszystko, co można zrobić za pomocą programu SPSS byłaby kilkanaście, a może nawet kilkaset razy bardziej obszerna. Proszę, aby Czytelnik nie obawiał się – opisane sposoby naprawdę wystarczą pedagogowi, a tylko w incydentalnych przypadkach będzie potrzebny lepszy podręcznik.

# Opracowania statystyczne

*Z tego rozdziału nauczysz się:*

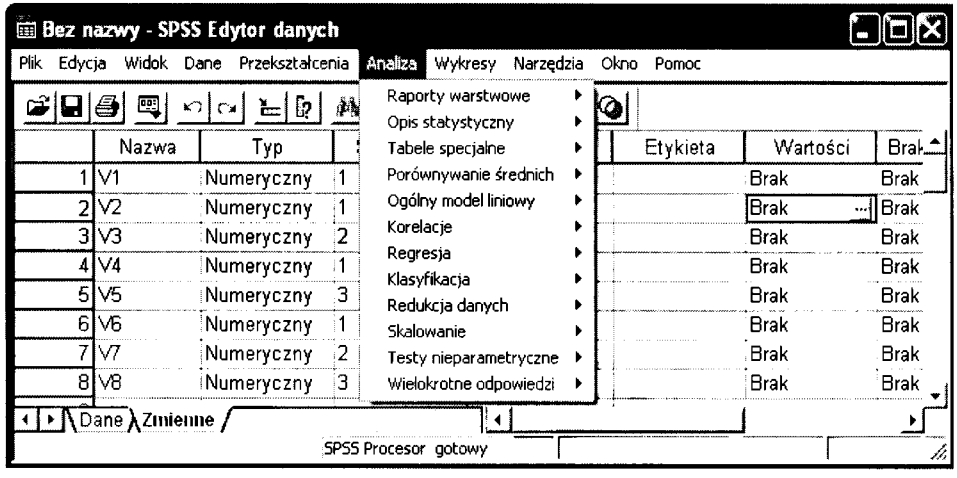
- *obliczać medianę i wartość modalną*
- *obliczać średnią arytmetyczną*
- *obliczać wariancję*
- *obliczać odchylenie standardowe*
- *robić tabele liczebności dla wybranej zmiennej*
- *robić tabele dwóch powiązanych zmiennych*
- *obliczać różne współczynniki korelacji*
- *robić chi-kwadrat test niezależności*
- *obliczać różne współczynniki zbieżności*
- *robić chi-kwadrat test zgodności*

# Wstęp

Opracowania statystyczne uruchamia się w oknie z siecią danych. Menu Analiza zawiera szereg podmenu i poszczególnych opcji.

Analyze

okno Data Editor




Rysunek 27. Menu Analiza

W pracach badawczych studentów pedagogiki i nauczycieli najbardziej przydatne są opcje wymienione w tabeli.

Tabela 12. Przydatne opracowania statystyczne

| polskie menu | angielskie menu |
|---|---|
| **Opis statystyczny** | **Descriptive Statistics** |
| ☐☐ Częstości | ☐☐ Frequencies |
| ☐☐ Statystyki opisowe | ☐☐ Descriptives |
| ☐☐ Tabele krzyżowe | ☐☐ Crosstabs |
| **Korelacje** | **Correlate** |
| ☐☐ Parami | ☐☐ Bivariate |
| ☐☐ Cząstkowe | ☐☐ Partial |
| **Testy nieparametryczne** | **Nonparametric Tests** |
| ☐☐ Chi-kwadrat | ☐☐ Chi-Square |

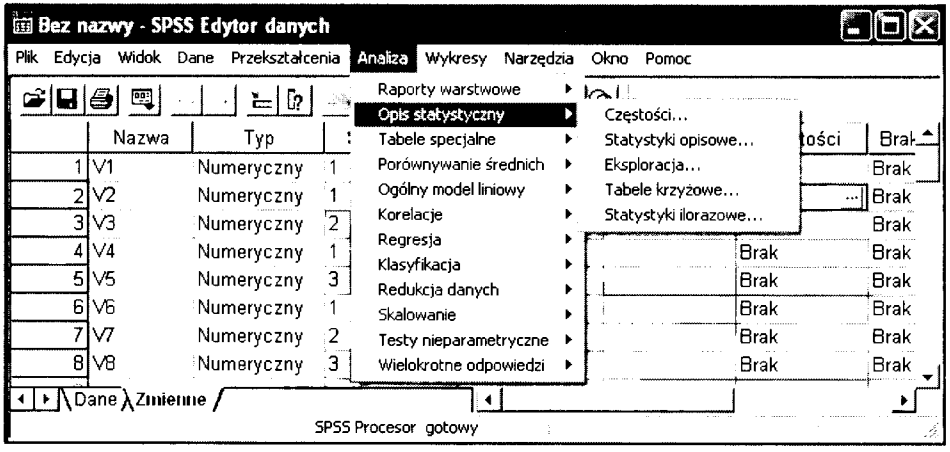
W każdym podmenu proponowane są różne opracowania (polecenia). Po wyborze polecenia otwiera się okno dialogowe do wykonania tego opracowania. Okna są do siebie bardzo podobne, dlatego każde z nich jest oznaczone nazwą opracowania (np. Częstości, Chi-kwadrat test itd.). W oknie dialogowym wybiera się różne szczegóły opracowania.

# Opis statystyczny

Descriptive Statistics

Z pięciu poleceń tego menu w badaniach edukacji wystarczą najczęściej trzy: Częstości, Statystyki opisowe i Tabele krzyżowe. Polecenie Częstości służy do prostych opracowań danych jakościowych, polecenie Statystyki opisowe do prostych opracowań danych ilościowych a polecenie Tabele krzyżowe do tworzenia dwuwymiarowych tabel i przeprowadzania testu Chi- kwadrat.

Frequencies
Descriptives
Crosstabs




Rysunek 28. Menu Opis statystyczny

## Częstości

Frequencies

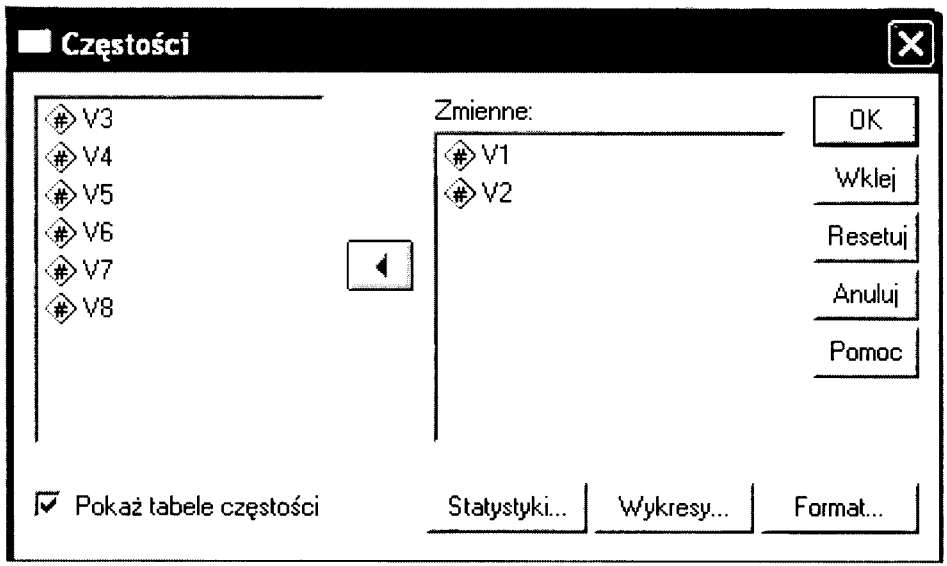
To program, który robi tabele liczebności wybranych zmiennych i wylicza niektóre wskaźniki.

Otwórz menu **Analiza**, wybierz podmenu **Opis statystyczny** i kliknij na pozycji **Częstości.** Otworzy się okno, w którym można wybrać wymagane szczegóły opracowania.

Analyze+
Descriptive
Statistics+
Frequencies

okno Frequencies

przenieś zmienne w prawe pole



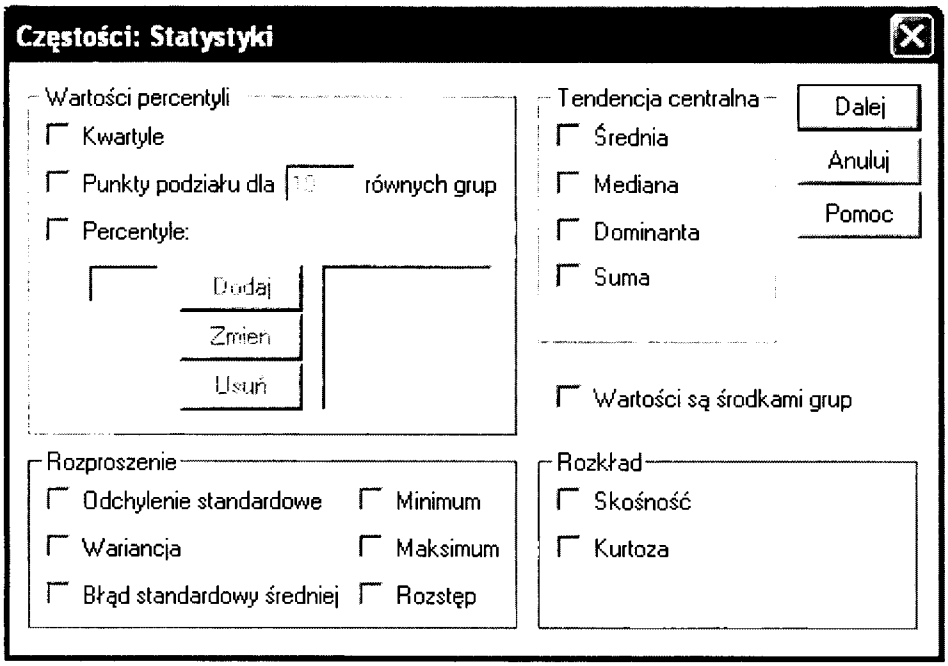

Rysunek 29. Okno Częstości

Frequencies

W lewym polu umieszczony jest spis wszystkich zmiennych, a w prawym te zmienne, które mają być opracowane za pomocą programu Częstości. Wybraną zmienną należy zaznaczyć w lewym polu i kliknąć na strzałkę pomiędzy lewym a prawym polem. Wybrana zmienna zostanie przeniesiona w prawe pole. W podobny sposób można wrócić zmienne z prawego pola na lewe. Strzałka automatycznie obraca się w lewo, jeżeli zaznaczona została zmienna w prawym polu. Zmienne można przenosić jedną po drugiej lub więcej jednocześnie. Do przenoszenia kilku zmiennych jednocześnie należy trzymać klawisz **CTRL** i zaznaczać zmienne. Gdy wszystkie zmienne zostaną zaznaczone należy puścić CTRL i kliknąć na strzałkę. Można wybrać także parametry, które program ma wyliczyć. Należy kliknąć na przycisk **Statystyki**. W nowym oknie trzeba zaznaczyć wszystkie żądane parametry (Średnia, Mediana, Wariancja, itd.).

Statistics

Mean
Median
Variance




okno Frequencies: Statistics

Rysunek 30. Okno Częstości: Statystyki

Tabela 13. Najbardziej potrzebne opcje opracowania Częstości

przydatne parametry

| spis parametrów | opis parametrów | angielska nazwa |
|---|---|---|
| Średnia | M (średnia arytmetyczna) | Mean |
| Mediana | Me | Median |
| Dominanta | Mo ( wartość modalna) | Mode |
| Suma | suma wszystkich wartości | Sum |
| Odchylenie standardowe | $\sigma$ | Std. deviation |
| Wariancja | $\sigma^2$ | Variance |
| Minimum | najniższa wartość | Minimum |
| Maksimum | najwyższa wartość | Maksimum |
| Rozstęp | przedział, w którym rozproszone są wszystkie rezultaty | Range |
| Kwartyle, Punkty podziału do kilku równych grup, Percentyle, Skośność, Kurtoza | rzadko używane wskaźniki | |

średnia
mediana

wariancja
minimum
maksimum
rozstęp

Continue
OK

Po zaznaczeniu opcji należy nacisnąć przycisk **Dalej** i okno Statystyki zamknie się. W głównym oknie Częstości kliknij **OK** i poczekaj kilka chwil, aby SPSS opracował dane.

Frequencies

Z możliwości wyliczania parametrów korzysta się rzadko, ponieważ zwykle opracowuje się programem Częstości dane jakościowe.

Charts

W podobny sposób można skorzystać też z opcji Wykresy. Po wybraniu polecenia **Wykresy** otworzy się okno Częstości: Wykresy i można dokonać wyboru z kilku proponowanych wykresów.


Rysunek 31. Okno Częstości: Wykresy

Descriptives

## *Statystyki opisowe*

Frequencies

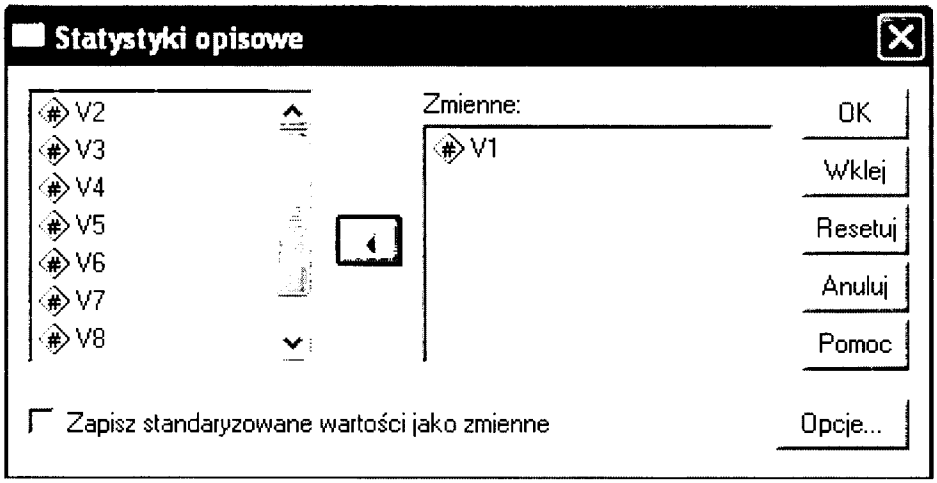
Program Statystyki opisowe jest bardzo podobny do programu Częstości. Używa się go do opracowywania danych ilościowych. Program wylicza podobne parametry, a główna różnica pomiędzy nim, a programem Częstości polega na tym, że program Statystyki opisowe nie tworzy

tabel. Tabele liczebności są bowiem w przypadku większego rozproszenia danych długie i nieprzejrzyste – a dane ilościowe są prawie zawsze bardzo rozproszone. W licznych przypadkach takie tabele zajmują kilka stron A4. Oprócz tego nie ma potrzeby robienia tabel, nawet w przypadku, gdy są one krótke. W przypadku danych jakościowych najczęściej interpretuje się tabele liczebności, a w przypadku danych ilościowych interpretuje się głównie parametry. Z tego powodu dla danych ilościowych lepszym rozwiązaniem jest stosowanie programu Statystyki opisowe, a dla danych jakościowych stosowanie programu Częstości. Ze względu na fakt, iż programy te są do siebie podobne, można w obu przypadkach bez wielkich strat zastosować ten drugi »niewłaściwy« program.

Descriptive statistics

Otwórz menu **Analiza**, wybierz podmenu **Opis statystyczny** i kliknij na pozycji **Statystyki opisowe**. Otworzy się główne okno programu, w którym wybiera się zmienne i szczegóły opracowania (w taki sam sposób, jak w oknie Częstości).

Analyze+ Descriptive statistics+ Descriptives

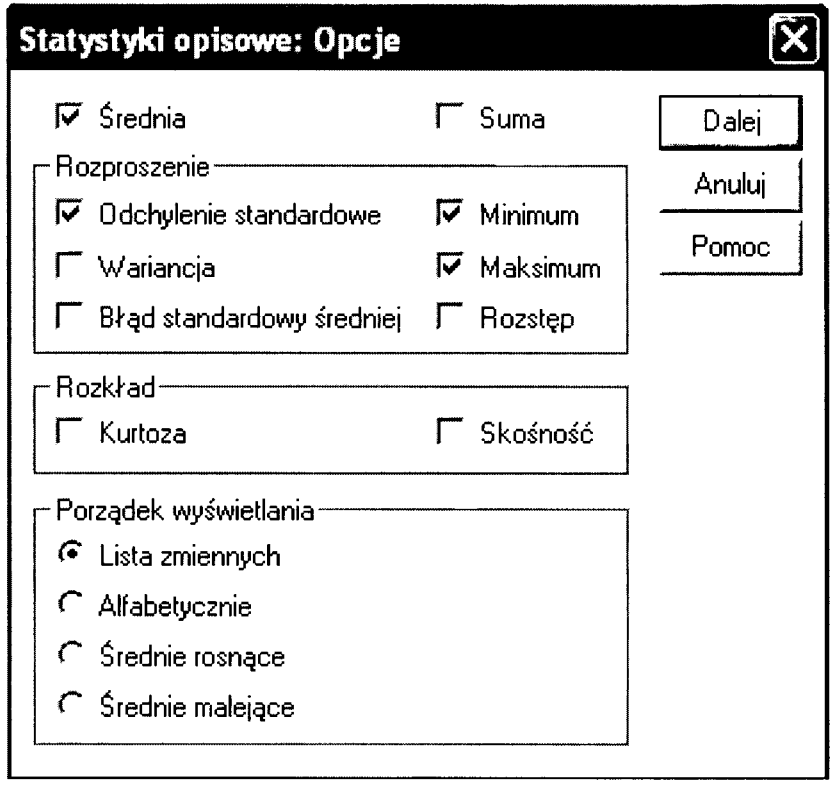



Rysunek 32. Okno Statystyki opisowe

Po wyborze zmiennych kliknij na przycisk **Opcje.** W nowym oknie wybiera się parametry.

Options

okno Descriptive statistics: Options




Rysunek 33. Okno Statystyki opisowe: Opcje

Oprócz parametrów w tym oknie można wybrać jeszcze sposób uporządkowania wyników.

Variable list

-Wybierając opcję Lista zmiennych, wyniki będą uporządkowane według kolejności zmiennych na liście.

Alphabetic

-Wybierając opcję Alfabetycznie, wyniki będą uporządkowane alfabetycznie według imion zmiennych.

Ascending means

Descending means

-Wybierając opcję Średnie rosnące, wyniki będą uporządkowane od zmiennej z najmniejszą średnią arytmetyczną do zmiennej z największą średnią arytmetyczną, a wybierając opcje Średnie malejące, wyniki będą uporządkowane odwrotnie.

Continue OK

Po wybraniu dodatkowych możliwości, kliknij **Dalej** i po powrocie do głównego okna programu (Statystyki opisowe) jeszcze **OK**. Po kilku sekundach otrzymuje się wyniki.

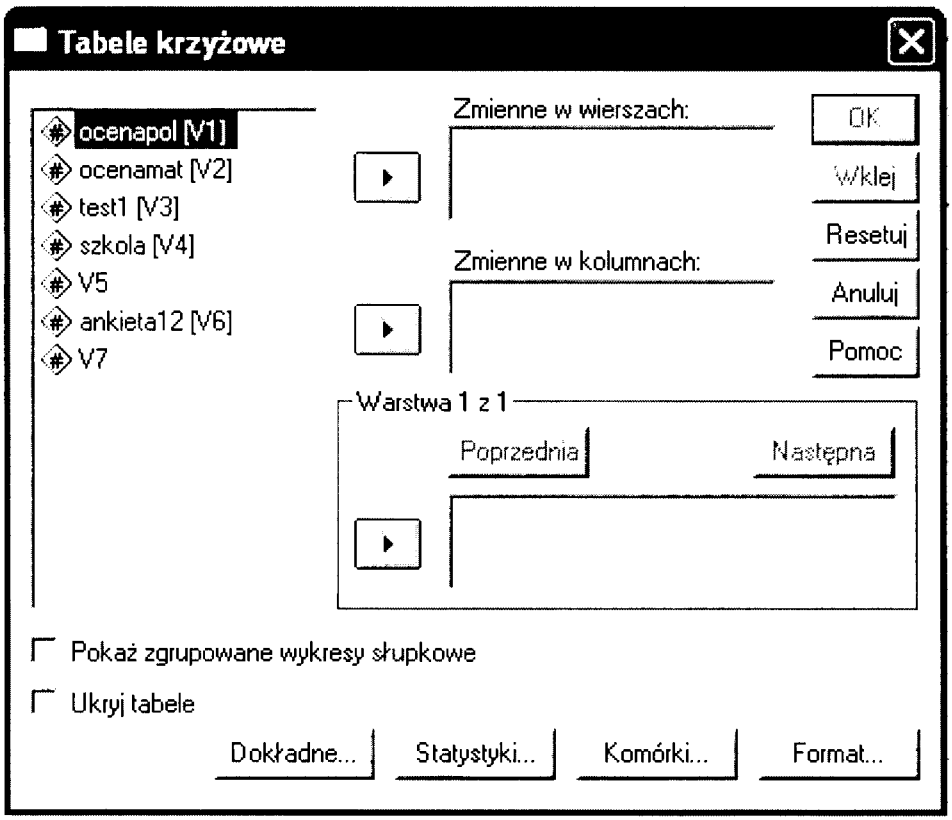
## Tabele krzyżowe

Crosstabs

Program ten zestawia tabele dwuwymiarowe, czyli te, które prezentują powiązanie dwóch zmiennych oraz wylicza wszystkie rezultaty do chi-kwadrat testu zależności.

Otwórz menu **Analiza**, wybierz podmenu **Opis statystyczny** i kliknij na pozycji **Tabele krzyżowe**. Otworzy się okno dialogowe.

Analyze+
Descriptive
Statistics+
Crosstabs




Row(s)

Column(s)

Rysunek 34. Okno Tabele krzyżowe

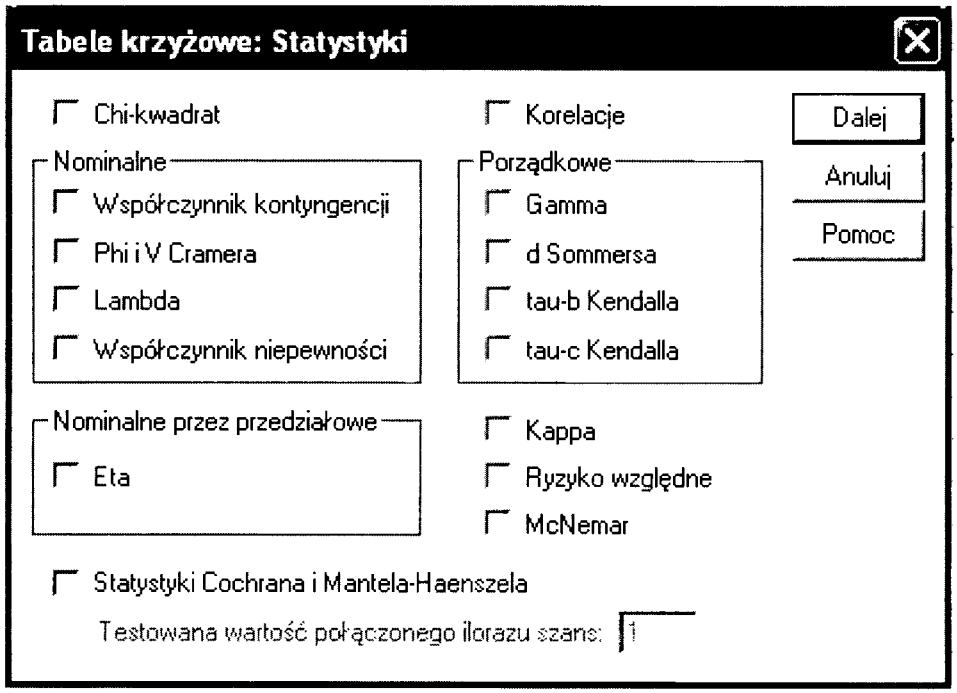
Z listy zmiennych w lewym polu należy przenieść w prawe górne pole (Zmienne w wierszach) jedną lub więcej zmiennych niezależnych. W prawe dolne pole (Zmienne w kolumnach) należy przenieść jedną lub więcej zmiennych zależnych. Program SPSS zestawi tabele dla każdej pary zmiennych: jednej niezależnej i jednej zależnej. Jeżeli w polu Zmienne w wierszach znajduje się pięć zmiennych, a w polu Zmienne

Row(s)
Column(s)

w kolumnach osiem, wtedy otrzymuje się 40 tabel. W wierszach tabel pojawią się kategorie zmiennej niezależnej, a w kolumnach kategorie zmiennej zależnej.

Statistics

Polecenie **Statystyki** w tym oknie umożliwia wyliczanie wartości chi-kwadrat i współczynników zbieżności (siły związku).




Rysunek 35. Okno Tabele krzyżowe: Statystyki

Tabela 14. Podstawowe wskaźniki opracowania Tabele krzyżowe

| spis parametrów | opis parametrów | angielska nazwa |
|---|---|---|
| Chi-kwadrat | wartość chi-kwadrat | Chi-square |
| Współczynnik kontyngencji | nieskorygowany współczynnik zbieżności Pearsona C | Contingency coefficient |
| Phi i V Cramera | współczynnik zbieżności Φ i współczynnik zbieżności Craméra | Phi and Cramér's V |
| Eta | stosunek korelacyjny | Eta |
| Korelacje | współczynnik korelacji Pearsona $r_{xy}$ | Correlations |

Chi square

Contingency coeficient

Kliknij na przycisk **Statystyki**, zaznacz wszystkie wyliczenia i kliknij na przycisk **Dalej.** Zamknie się okno Statystyki. Po powrocie do głównego okna Tabele krzyżowe można skorzystać też z opcji Komórki.

Statistics
Continue


Percentages

Rysunek 36. Okno Tabele krzyżowe: Zawartość komórek

W tym oknie wybiera się sposób wyliczania procentów w tabeli. Zwykle wybiera się wyliczanie odrębne dla każdego nowego wiersza (Procenty: W wierszu). Procenty w tabeli w tym przypadku pokazują wpływ zmiennej niezależnej na zmienną zależną. Problem ten został szczegółowo omówiony wcześniej.

Percentages

Kliknij na **Dalej,** a po powrocie do okna Tabele krzyżowe kliknij przycisk **OK.** Tym przyciskiem uruchamia się program Tabele krzyżowe. Po kilku sekundach otrzymuje się wyniki.

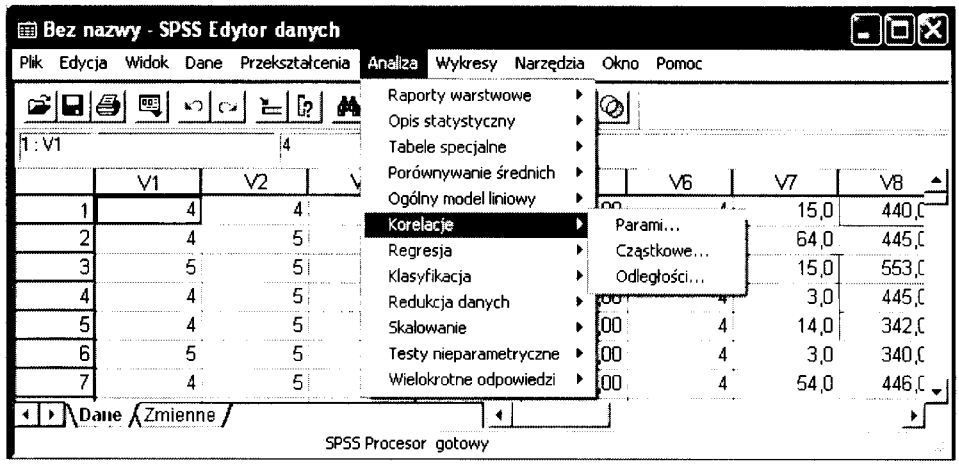
Continue
OK
Crosstabs

Correlate

# Korelacje

Bivariate
Partial

Do obliczania zwyczajnych współczynników korelacji Pearsona należy wybrać podprogram Parami, a do obliczania współczynników korelacji cząstkowej podprogram Cząstkowe.




Rysunek 37. Menu Korelacje

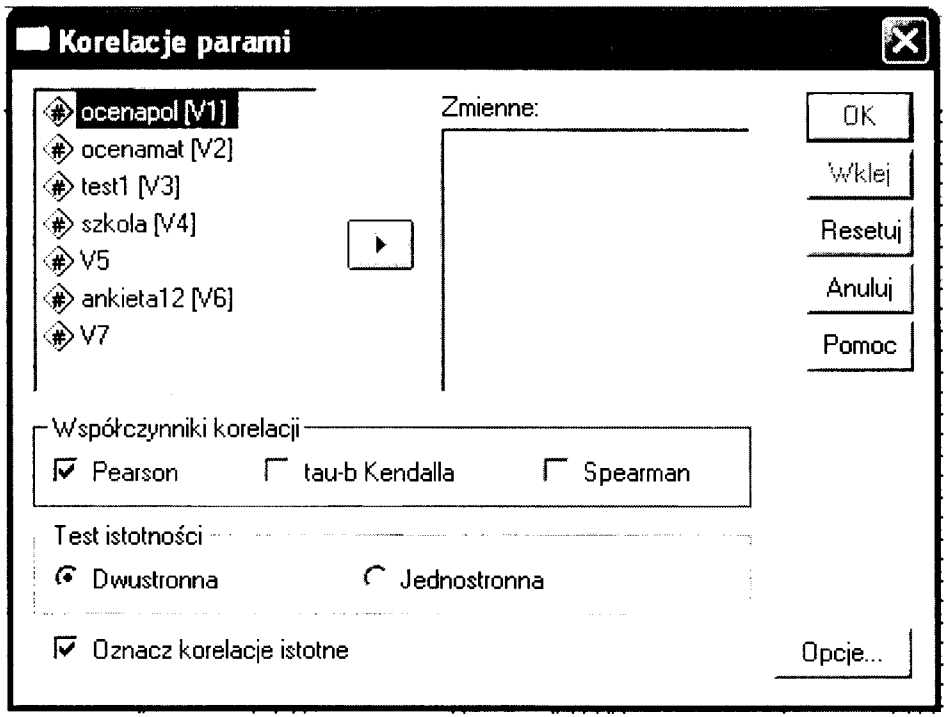
Bivariate
Correlation

## *Korelacje parami*

Program Parami wylicza współczynniki korelacji Pearsona i Spearmana oraz ustala ich poziom istotności (test hipotezy zerowej o współczynniku korelacji w populacji generalnej).

Analyze+
Correlate+
Bivariate

Otwórz menu **Analiza**, wybierz podmenu **Korelacje** i kliknij na pozycji **Parami.** Otworzy się okno dialogowe Korelacje parami.




okno Bivariate Correlations

Rysunek 38. Okno Korelacje parami

Z lewej listy na prawą (Zmienne) przenieś wszystkie zmienne, dla których mają być obliczone współczynniki korelacji. Program SPSS oblicza korelacje każdej zmiennej z każdą zmienną (z prawej listy). Wybierz także współczynniki i test hipotezy zerowej. W przypadku zmiennych przedziałowych lub ilorazowych (np. wiek, liczba godzin uczenia się, pensje itd.) wybiera się współczynnik Pearsona (rubryka Pearson). Jeżeli dane są w postaci rang (każdy uczeń posiada swoją rangę – podobnie jak wyniki sportowe w biegu), wybiera się współczynnik Spearmana (rubryka Spearman). Można też wybrać tau-b Kendalla.

Variables

Pearson

Spearman
Kendall's tau-b

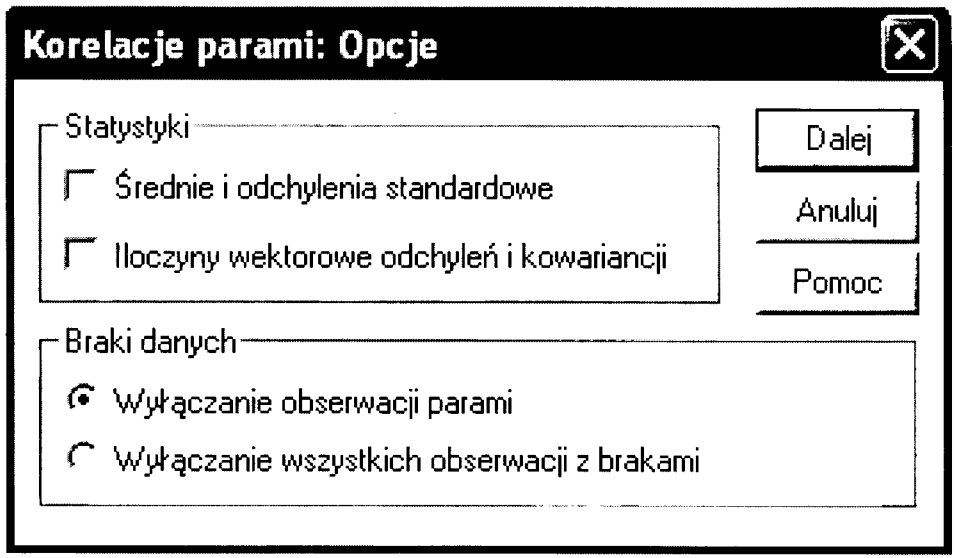
Do testowania hipotezy zerowej o współczynnikach korelacji stosuje się zwykle dwustronny test istotności, a test jednostronny stosuje się rzadko.

Test of Significance

Oznakuj jeszcze polecenie na dnie okna: Oznacz korelacje istotne. To oznacza, że istotne współczynniki korelacji będą oznaczone gwiazdką. Ułatwia to przegląd wyników.

Flag significant correlations

okno Bivariate Correlation: Options




Rysunek 39. Okno Korelacje parami: Opcje

Continue

Bivariate Correlation OK

Z okna Opcje do okna głównego wraca się przyciskiem na **Dalej**. Gdy wszystko zostało wybrane i oznaczone, kliknij w głównym oknie Korelacje parami przycisk **OK.** Za kilka chwil otrzymuje się wyniki.

Program Korelacje parami ma dość istotną wadę: wylicza korelacje każdej zmiennej z każdą zmienną na liście w prawym polu (nawet też samej z sobą – zupełnie niepotrzebnie). Jeżeli wybierze się za każdym razem po dwie zmienne, to oznacza, że konieczne będzie powtarzanie tej samej procedury tyle razy, ile jest par zmiennych. Jeżeli wybierze się wszystkie zmienne, oznacza to, iż w końcu z tabeli wyników trzeba będzie usuwać więcej niż połowę wszystkich współczynników. Zostanie to zilustrowane przykładem. Chcemy wyliczyć korelacje pomiędzy wzrostem (V1) a dziesięcioma innymi zmiennymi (celem jest otrzymanie dziesięciu współczynników korelacji). Do listy wpiszemy wszystkie jedenaście zmiennych. SPSS wylicza korelacje każdej z każdą, dlatego też otrzymuje się tabele zawierającą 121 współczynników korelacji (11x11). Potrzebujemy ich tylko dziesięć, należy zatem z tabeli usunąć 111 niepotrzebnych współczynników. Oto przykład otrzymanej tabeli:

Tabela 15. Współczynniki korelacji Pearsona

| | V1 | V2 | V3 | V4 | V5 | V6 | V7 | V8 | V9 | V10 | V11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| V1 | 1 | ,57 | ,67 | -,12 | ,42 | ,34 | ,47 | ,41 | ,45 | ,02 | ,38 |
| V2 | ,57 | 1 | ,66 | -,10 | ,31 | ,52 | ,36 | ,53 | ,45 | ,05 | ,32 |
| V3 | ,67 | ,66 | 1 | -,59 | ,36 | ,40 | ,43 | ,45 | ,46 | ,04 | ,35 |
| V4 | -,12 | -,10 | -,59 | 1 | -,09 | -,08 | -,09 | -,07 | -,07 | -,04 | -,52 |
| V5 | ,42 | ,31 | ,36 | -,09 | 1 | ,37 | ,43 | ,36 | ,37 | ,12 | ,58 |
| V6 | ,34 | ,52 | ,40 | -,08 | ,37 | 1 | ,33 | ,47 | ,37 | ,36 | ,32 |
| V7 | ,47 | ,36 | ,43 | -,09 | ,43 | ,33 | 1 | ,51 | ,61 | ,15 | ,60 |
| V8 | ,41 | ,53 | ,44 | -,07 | ,36 | ,47 | ,51 | 1 | ,64 | ,20 | ,49 |
| V9 | ,45 | ,45 | ,46 | -,07 | ,37 | ,37 | ,61 | ,64 | 1 | -,43 | ,54 |
| V10 | ,02 | ,05 | ,04 | -,04 | ,12 | ,36 | ,15 | ,20 | -,43 | 1 | -,04 |
| V11 | ,38 | ,32 | ,35 | -,52 | ,58 | ,32 | ,60 | ,49 | ,54 | -,04 | 1 |

Druga możliwość polega na tym, że dziesięć razy wracamy do punktu wyjścia i ponownie uruchamiamy program Korelacje parami, za każdym razem wybierając wzrost i jedną z pozostałych zmiennych.

Correlation Bivariate

## *Korelacje cząstkowe*

Partial Correlations

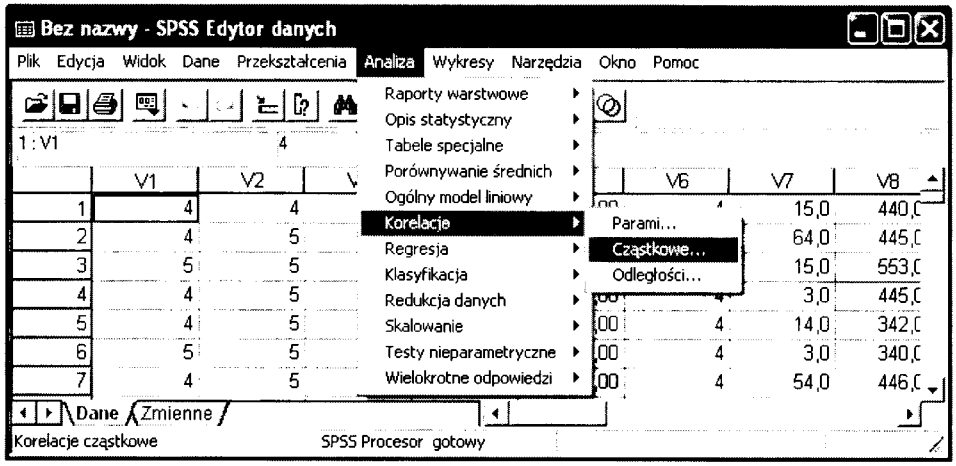
Program wylicza współczynnik korelacji cząstkowej Pearsona oraz ustala jego poziom istotności (test hipotezy zerowej o współczynniku korelacji w populacji generalnej). Mankamentem zwyczajnych współczynników korelacji jest fakt, że każda zmienna, która wpływa na *x* zniekształca obraz korelacji między *x* i *y*. Z tego powodu zwyczajny współczynnik korelacji zawiera oprócz głównego wpływu (*x* na *y*), także uboczne wpływy innych zmiennych. Oznacza to, że nie ma pewności, iż współczynnik korelacji odzwierciedla rzeczywiście jedynie wpływ *x* na *y*. Współczynnik ten nie jest „czystym". Uboczne wpływy należy wyeliminować, najlepiej za pomocą procedury statystycznej. Rezultatem eliminacji wpływów

ubocznych jest korelacja cząstkowa. Eliminacja jednego wpływu ubocznego jest najprostszym przypadkiem. Współczynnik korelacji cząstkowej pokazuje wówczas współzależność między zmiennymi *x* i *y*, bez tego wpływu ubocznego. Można eliminować też wpływ większej liczby zmiennych, wymaga to jednak od badacza pomiaru wszystkich tych zmiennych i bardziej złożonego wyliczenia współczynników. Przy korzystaniu z SPSS badaczowi zostaje już tylko pierwsza cześć tego zadania, ponieważ skomplikowane obliczenia wykonuje komputer.

Analyze+
Correlate+
Partial

Otwórz menu **Analiza**, wybierz podmenu **Korelacje** i kliknij na pozycji **Cząstkowe**.

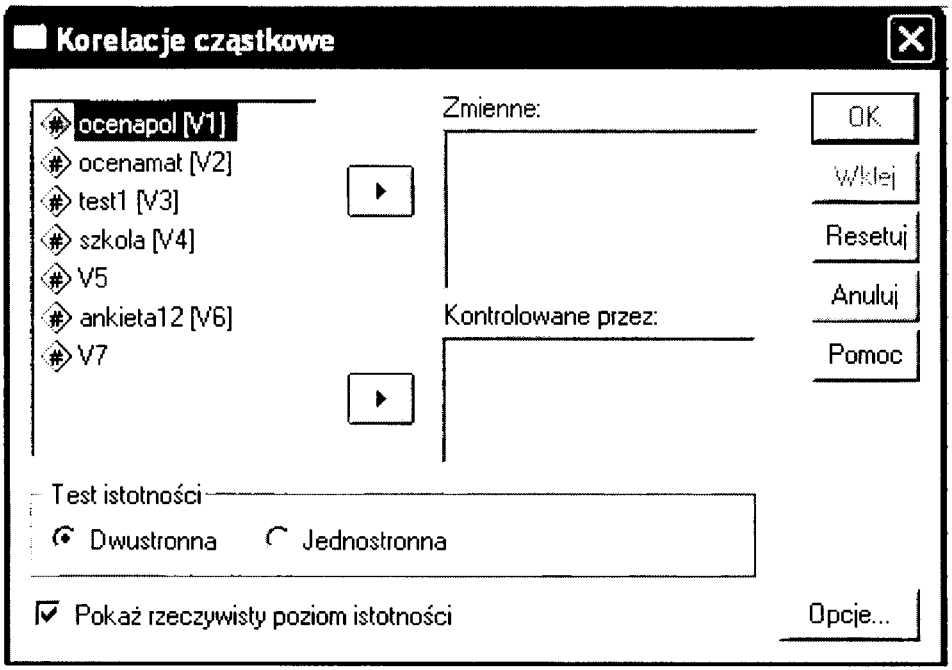
Partial
Correlations




Rysunek 40. Menu Korelacje cząstkowe

Partial
Correlations

Otworzy się okno dialogowe Korelacje cząstkowe. Z powodu przejrzystości opisu procedury, przykład przedstawiony w dalszej części dotyczy obliczania korelacji tylko dla jednej pary zmiennych (niezależna *x* i zależna *y*). SPSS umożliwia oczywiście także wybór większej liczby zmiennych.




okno Partial Correlations

Rysunek 41. Okno Korelacje cząstkowe

Przenieś z lewej listy w prawe górne pole (Zmienne) te dwie zmienne, dla których ma być obliczony współczynnik korelacji cząstkowej. W prawe dolne pole (Kontrolowane przez) przenieś jedną lub więcej zmiennych, których wpływ należy wyeliminować. Jeżeli w dolnym polu znajduje się tylko jedna zmienna – otrzymuje się współczynnik korelacji cząstkowej pierwszego rzędu, przy dwóch zmiennych – współczynnik korelacji cząstkowej drugiego rzędu, itd.

Variables

Controlling for

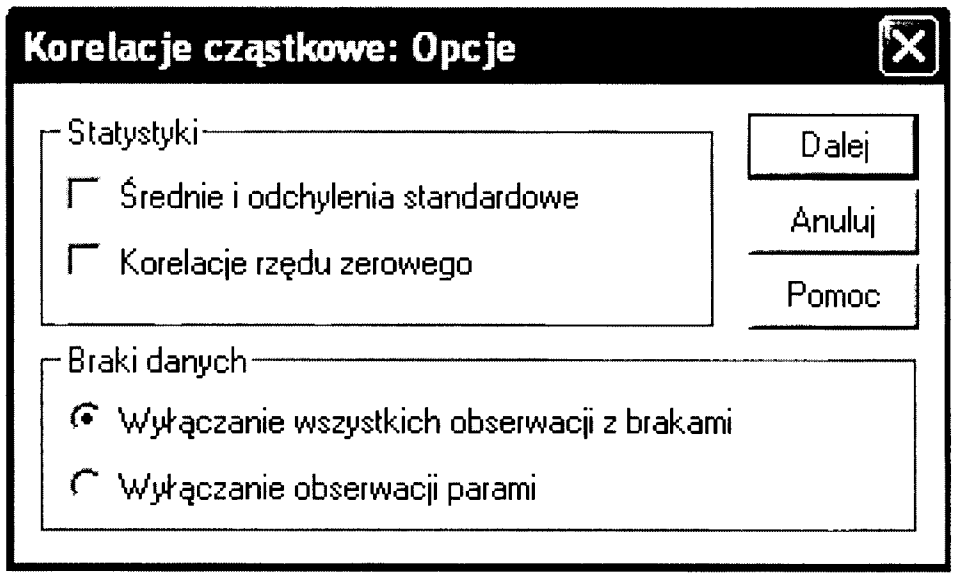
Wybierz też test hipotezy zerowej. Do testowania hipotezy zerowej o współczynnikach korelacji stosuje się zwykle test dwustronny, a test jednostronny stosuje się rzadko.

Test of Significance
Two-tailed
One-tailed

Można też wybrać dodatkowe wyliczenia. Służy do tego przycisk **Opcje**.

Options

okno Partial Correlations: Options




Rysunek 42. Okno Korelacje cząstkowe: Opcje

Zero-order correlations

W tym oknie przydatną jest opcja obliczania zwykłego współczynnika korelacji (bez usunięcia ubocznych wpływów). Może on posłużyć jako dobra kontrola stanu wyjściowego. Aby go otrzymać, należy oznaczyć drugie polecenie od góry (Korelacje rzędu zerowego).

Partial Correlations OK

Gdy wszystko zostało wybrane i oznaczone, kliknij w głównym oknie Korelacje cząstkowe przycisk **OK.** Za kilka chwil otrzymuje się wyniki.

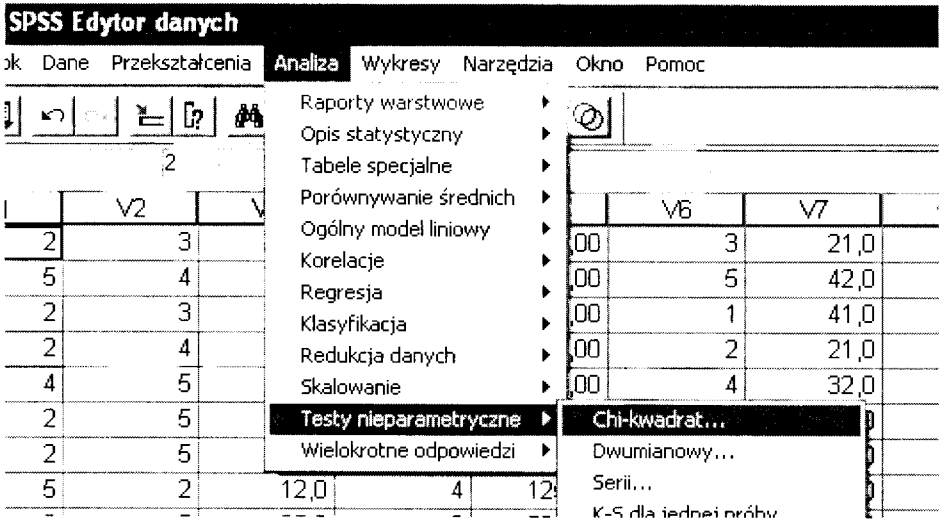
Nonparametric Tests

## Testy nieparametryczne

W tym programie najczęściej wybiera się test zgodności. Test ten przeprowadza się wówczas, gdy znane są dane z próby i dla wybranej zmiennej sprawdza się, jaki jest rozkład kategorii tej zmiennej w populacji generalnej.

Analyze+ Nonparametric Tests+ Chi-Square

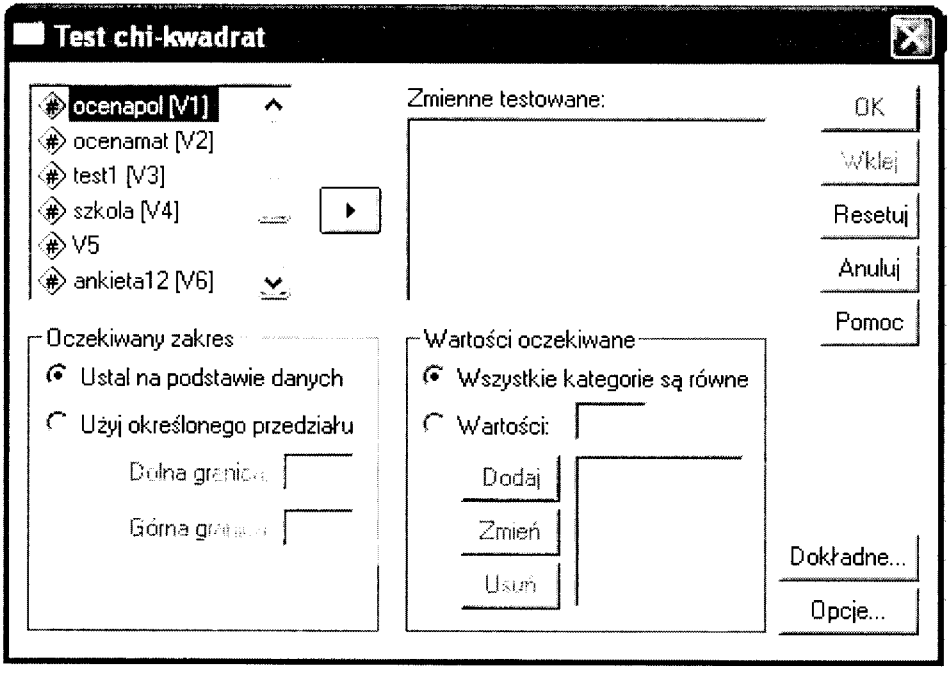
Otwórz menu **Analiza**, wybierz podmenu **Testy nieparametryczne** i kliknij na pozycji **Chi-kwadrat**.




okno Data Editor

Rysunek 43. Menu Testy nieparametryczne

Otworzy się okno tego podprogramu:



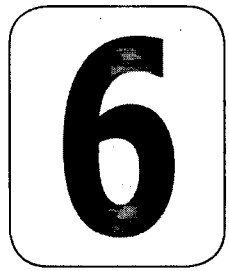

okno Chi-Square Test

Rysunek 44. Okno Test chi-kwadrat

Test Variable List
Expected Value
All categories equal
OK

Wybrane zmienne z lewej listy przenieś do prawego pola (Zmienne testowane). Automatycznie w polu Wartości oczekiwane wyświetla się warunek zerowej hipotezy: Wszystkie kategorie są równe.

Kliknij na **OK** i poczekaj na wynik.



# Wyniki

*Z tego rozdziału nauczysz się:*

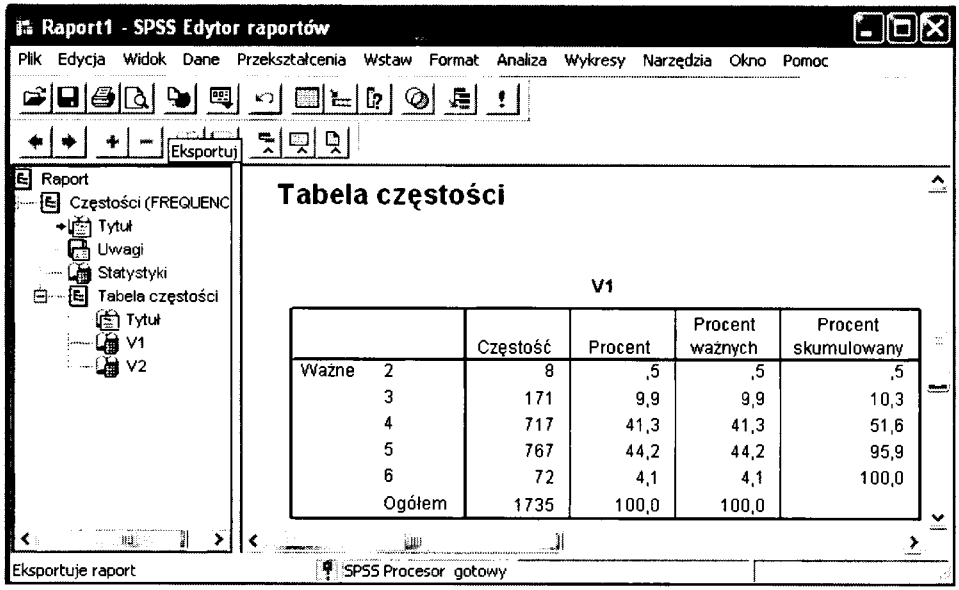
- *odczytywać wyniki z opracowań*
- *odczytywać zestawy tabelaryczne*
- *poszukiwać potrzebne rezultaty*
- *usuwać niepotrzebne rezultaty*
- *przygotowywać dane do opracowania Tabele krzyżowe*

Output Viewer

# Edytor raportów

Po uruchomieniu każdego opracowania otworzy się okno Edytor raportów:

okno Output Viewer




Rysunek 45. Okno Edytor raportów

W lewym polu okna znajduje się spis wszystkich rezultatów. Po kliknięciu na wybrany element ze spisu, pojawi się on w prawym polu (tabela, wykres itd.).

Frequencies

# Program Częstości

Wykres pokazuje rozkład wyników i procenty dla zmiennej płeć.

rozkład wyników

Tabela 16. Rozkład ważnych i brakujących odpowiedzi

| | | |
|---|---|---|
| N | Ważne | 1735 |
| | Braki danych | 0 |

Tabela 17. Odpowiedzi dla zmiennej płeć

| | | Częstość | Procent | Procent ważnych | Procent skumulowany |
|---|---|---|---|---|---|
| Ważne | 1 | 896 | 51,6 | 51,6 | 51,6 |
| | 2 | 839 | 48,4 | 48,4 | 100,0 |
| | Ogółem | 1735 | 100,0 | 100,0 | |

tabela zawiera jedną zmienną

Kategoria 1 (np. kobiety) posiada liczebność 896 lub 51,6%, a kategoria 2 (np. mężczyźni) – liczebność 839 lub 48,4% z liczebności całej próby n=1735. Procenty w trzeciej i czwartej kolumnie są identyczne, ponieważ nie występują brakujące dane. Podobnie jest we wszystkich tabelach, w których nie występuje brak danych.

Poniżej zaprezentowano tabelę nr 18 dla zmiennej ankieta1. Od tabeli nr 17 różni się liczbą kategorii.

Tabela 18. Odpowiedzi dla zmiennej ankieta1

| | | Częstość | Procent | Procent ważnych | Procent skumulowany |
|---|---|---|---|---|---|
| Ważne | 2 | 137 | 7,9 | 7,9 | 7,9 |
| | 3 | 539 | 31,1 | 31,1 | 39,0 |
| | 4 | 698 | 40,2 | 40,2 | 79,2 |
| | 5 | 316 | 18,2 | 18,2 | 97,4 |
| | 6 | 45 | 2,6 | 2,6 | 100,0 |
| | Ogółem | 1735 | 100,0 | 100,0 | |

tabela zawiera jedną zmienną

W obydwu tabelach z powodu przejrzystości zostawiono kategorie numeryczne (jak w pliku danych). Pomimo iż istniała też możliwość odpowiedzi 1, nikt z respondentów jej nie wybrał i dlatego nie występuje ona w tabeli. Nie pojawia się też brak odpowiedzi, stąd trzecia i czwarta kolumna są identyczne. Kolejna tabela zawiera odpowiedzi dla zmiennej ocena z chemii.

kategorie numeryczne

Tabela 19. Rozkład ważnych i brakujących odpowiedzi

| | | |
|---|---|---|
| N | Ważne | 1670 |
| | Braki danych | 65 |

Tabela 20. Odpowiedzi dla zmiennej ocena z chemii

| | | Częstość | Procent | Procent ważnych | Procent skumulowany |
|---|---|---|---|---|---|
| Ważne | 1 | 3 | 0,2 | 0,2 | 0,2 |
| | 2 | 157 | 9,0 | 9,4 | 9,6 |
| | 3 | 553 | 31,9 | 33,1 | 42,7 |
| | 4 | 649 | 37,4 | 38,9 | 81,6 |
| | 5 | 292 | 16,8 | 17,5 | 99,0 |
| | 6 | 16 | 0,9 | 1,0 | 100,0 |
| | Ogółem | 1670 | 96,3 | 100,0 | |
| Braki danych | 0 | 65 | 3,7 | | |
| Ogółem | | 1735 | 100,0 | | |

tabela zawiera jedną zmienną

oceny uczniów

Ocenę negatywną (1) mają trzej uczniowie, ocenę dopuszczającą (2) otrzymało 157 uczniów, ocenę dostateczną (3) uzyskało 553 uczniów, ocenę dobrą (4) – 649 uczniów, ocenę bardzo dobrą (5) 292 i ocenę celującą 16 uczniów. Liczba ocenionych uczniów wynosi 1670. Z tej liczebności wyliczono procenty w czwartej kolumnie (kolumna Procent ważnych). Na przykład 553 uczniów z oceną dostateczną stanowi 33,1% wszystkich ocenionych uczniów. W próbie jest także 65 uczniów nieocenionych lub brak informacji o ich ocenach. Procenty w trzeciej kolumnie odnoszą się do liczby wszystkich uczniów w próbie – tych z ocenami i bez ocen (kolumna Procent). Tych samych 553 uczniów z oceną dostateczną stanowi 31,9% wszystkich uczniów w próbie. Czasami w badaniach istnieje potrzeba wyliczania wartości procentowych z całej liczebności, a niekiedy jedynie tylko z liczby wszystkich ważnych odpowiedzi.

brakujące dane

liczba ważnych odpowiedzi

liczba wszystkich respondentów w próbie

Po przeniesieniu tabeli do tekstu i jej zinterpretowaniu (np. w pracy magisterskiej, dysertacji, raporcie z badań itd.), niepotrzebną kolumnę należy usunąć. Najczęściej usuwa się też ostatnią kolumnę z liczebnościami skumulowanymi.

Frequencies

Tabela na kolejnej stronie zawiera przykład obliczonych parametrów,

które proponuje program Częstości.

Tabela 21. Wartości parametrów

| N | Ważne | 1735 |
|---|---|---|
| | Braki danych | 0 |
| Średnia | | 38,3816 |
| Mediana | | 38,0000 |
| Dominanta | | 38,00 |
| Odchylenie standardowe | | 5,79610 |
| Wariancja | | 33,595 |
| Rozstęp | | 38,00 |
| Minimum | | 18,00 |
| Maksimum | | 56,00 |
| Suma | | 66592,00 |
| Percentyle | 25 | 35,0000 |
| | 50 | 38,0000 |
| | 75 | 42,0000 |

wyliczone parametry

Descriptive Statistics

# Program Statystyki opisowe

Raport o wykonaniu opracowania Statystyki opisowe jest bardzo krótki i przejrzysty. Zawiera tylko jedną tabelę dla wszystkich wybranych zmiennych, w której umieszczone są parametry, brak natomiast tabel liczebności. Poniższa tabela jest przykładem takiego raportu.

Tabela 22. Wartości parametrów dla trzech zmiennych

| | N | Minimum | Maksimum | Średnia | Odchylenie standardowe | Wariancja |
|---|---|---|---|---|---|---|
| wiek | 1735 | 28,00 | 59,00 | 46,11 | 4,89 | 23,91 |
| test1 | 1735 | 22,00 | 72,00 | 47,06 | 8,66 | 75,02 |
| test2 | 1735 | 11,00 | 81,00 | 46,63 | 12,38 | 153,26 |

Minimum
Maximum
Mean
Standard deviation
Variance

Dla zmiennych: wiek i rezultaty dwóch testów wyliczono najniższą oraz

Oto krótka interpretacja rezultatów:

N of Sample

- liczebność próby wynosi 1735,
- najmłodszy respondent ma 28 lat, a najstarszy 59 lat, przeciętny wiek wynosi niewiele ponad 46 lat,
- odchylenie standardowe wynosi 4,89, czyli wiek respondentów różni się w przybliżeniu o pięć lat w górę i w dół od średniej arytmetycznej,
- wariancja wynosi 23,91,
- średnie arytmetyczne obydwu testów są prawie identyczne, natomiast rozproszenie rezultatów różni się zasadniczo (rezultaty pierwszego testu wahają się od 22 do 72, drugiego od 11 do 81),
- odchylenie standardowe i wariancja rezultatów drugiego testu są dużo większe.

Crosstabs

## Program Tabele krzyżowe

Wyniki rozpoczynają się informacją o liczebności respondentów, którzy udzielili odpowiedzi na obydwa pytania (Ważne) oraz o liczebności respondentów, którzy nie odpowiedzieli na obydwa pytania (Braki danych).

Tabela 23. Liczebności respondentów

Valid Cases+ Missing Cases Total

| | Obserwacje | | | | | |
|---|---|---|---|---|---|---|
| | Uwzględnione | | Wykluczone | | Ogółem | |
| | N | Procent | N | Procent | N | Procent |
| V1 * V2 | 1735 | 100,0% | 0 | ,0% | 1735 | 100,0% |
| V1 * V4 | 1735 | 100,0% | 0 | ,0% | 1735 | 100,0% |
| V5 * V2 | 1735 | 100,0% | 0 | ,0% | 1735 | 100,0% |
| V5 * V4 | 1735 | 100,0% | 0 | ,0% | 1735 | 100,0% |
| V5 * V11 | 1670 | 96,3% | 65 | 3,7% | 1735 | 100,0% |

Missing Cases

Crosstabs

Do tabeli nr 23 nie mogą być włączone rezultaty respondentów, którzy udzielili tylko jedną odpowiedź lub nie udzielili żadnej. Z tego powodu liczebność próby w poszczególnych tabelach różni się. Jeżeli brakujące odpowiedzi zostały wpisywane w postaci zera, należy przed opracowaniem Tabele krzyżowe dla wszystkich takich zmiennych określić zero jako brak odpowiedzi. W przeciwnym wypadku zero występuje w roli zwyczajnej kategorii tych zmiennych – co oznacza, że zostaną też uwzględnieni respondenci, którzy nie udzielili odpowiedzi. Jest to jednak zła sytuacja, ponieważ kategoria zero w tych przypadkach jest kategorią „bez treści" (nic nie oznacza i nie można tej „kategorii" zinterpretować merytorycznie). Zupełnie inna sytuacja występuje w przypadku, gdy zero pojawi się np. w pytaniu: "Ile papierosów dziennie Pani/Pan pali?". Tutaj zero jest odpowiedzią sensowną; takie zero zasadniczo różni się od braku odpowiedzi. Jeżeli zero jest odpowiedzią sensowną, należy dla braku odpowiedzi wybrać inny znak (spację lub coś innego). Wszystkich sytuacji w opracowywaniu badań i tak nie da się przewidzieć, dlatego też lepiej je dobrze rozumieć, aby samodzielnie znaleźć klucz do ich rozwiązania, niż posługiwać się tylko gotowymi receptami niektórych rozwiązań.

Crosstabs
Count
Percent

Drugą tabelą w raporcie programu Tabele krzyżowe jest tabela liczebności i procentów kategorii. Jeżeli do opracowania Tabele krzyżowe wybrano kilka par zmiennych, w raporcie znajdzie się oczywiście kilka tabel liczebności. Poniższa tabela zawiera odpowiedzi dla zmiennych: płeć i ankieta1.

Tabela 24. Odpowiedzi dla zmiennych płeć i ankieta1

tabela zawiera dwie zmienne

| | | ankieta1 | | | | | Ogółem |
|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | |
| płeć | 1 | 14 | 210 | 440 | 220 | 12 | 896 |
| | | 1,6% | 23,4% | 49,1% | 24,6% | 1,3% | 100,0% |
| | 2 | 0 | 22 | 193 | 537 | 87 | 839 |
| | | ,0% | 2,6% | 23,0% | 64,0% | 10,4% | 100,0% |
| Ogółem | | 14 | 232 | 633 | 757 | 99 | 1735 |
| | | ,8% | 13,4% | 36,5% | 43,6% | 5,7% | 100,0% |

W tabeli nr 24 płeć występuje w roli zmiennej niezależnej, a odpowiedzi na pierwsze pytanie ankiety w roli zmiennej zależnej. Procenty w tabeli

zostały wyliczone według kategorii płci (czyli zmiennej niezależnej).

Poniżej zostanie zaprezentowana techniczna część interpretacji procentów z tabeli nr 24 (1 oznacza kobiety, a 2 oznacza mężczyzn):

-odpowiedź 1 nie została wybrana przez respondentów,

-odpowiedź 2 zakreśliło 14 kobiet i nikt z mężczyzn (1,6% i 0,0%),

-odpowiedź 3 na pytanie ankietowe (ankieta1) wybrało więcej kobiet niż mężczyzn (23,4% kobiet i 2,6% mężczyzn),

-odpowiedź 4 została wybrana przez kobiety dużo częściej niż przez mężczyzn (49,1% kobiet i 23,0% mężczyzn),

-odpowiedź 5 częściej zakreślili mężczyźni (64,0% mężczyzn i 24,6% kobiet), podobnie też w odpowiedzi 6 (10,4% i 1,3%).

Do głębszej interpretacji jest potrzebna głębsza wiedza socjologiczna, dotycząca treści ankiety i wpływu płci na poglądy, a to nie jest przedmiotem niniejszej książki.

Chi-Square Test

Po każdej tabeli liczebności następuje tabela testu niezależności Chi- kwadrat (w przypadkach, jeżeli test został wybrany).

Tabela 25. Wyniki testu chi-kwadrat

| | Wartość | df | Istotność asymptotyczna (dwustronna) |
|---|---|---|---|
| Chi-kwadrat Pearsona | 450,90 | 4 | ,000 |
| N Ważnych obserwacji | 1735 | | |
| 0,0% komórek (0) ma liczebność oczekiwaną mniejszą niż 5. Minimalna liczebność oczekiwana wynosi 6,77. | | | |

Pearson Chi-Square
N of Valid Cases
expected counts less than 5

Tabela nr 25 pokazuje, że dla wybranej pary dwóch zmiennych wartość chi kwadrat wynosi 450,90. Poziom istotności dla g=4 jest bardzo wysoki (0,000). Hipotezę zerową (nie występuje zależność...) można odrzucić z ryzykiem mniejszym od 0,0% (aby obliczyć ryzyko należy pomnożyć poziom istotności przez 100%). Jeżeli stosuje się klasyczny sposób określania ryzyka za pomocą tabeli (trzy "okrągłe" wartości: 5%, 1% i 0,1%), należy napisać: "hipotezę zerową odrzuca się z ryzykiem mniejszym od 0,1%".

Significance

Pod tabelą jest umieszczona informacja dotycząca liczebności oczekiwanych. Na pierwszym miejscu znajduje się liczba i procent okienek z liczebnością mniejszą od 5, a na drugim najniższa oczekiwana liczebność. Do stosowania chi-kwadrat testu wszystkie oczekiwane liczebności powinny być większe od 5. Praktyka dowodzi jednak, że wymóg ten jest zbyt surowy. Toleruje się przypadki, w których procent liczebności niższych od 5 nie przekracza 20% i najniższa liczebność nie jest niższa od 1.

expected counts

W przypadku liczebności niższych od 5 najczęściej łączy się sąsiednie kategorie w tabeli (opisane wcześniej).

Uwaga: łączenie kategorii zmniejsza tabelę i liczbę stopni wolności. Program SPSS robi wszystko automatycznie, należy jednak zachować uwagę przy stosowaniu oryginalnych i zmniejszonych tabel.

degrees of freedom

Raport zawiera tabelę współczynników zbieżności (jeżeli ta opcja została wybrana).

Tabela 26. Współczynniki zbieżności

| | | Wartość | Istotność przybliżona |
|---|---|---|---|
| Nominalna przez Nominalna | V Kramera | ,51 | ,000 |
| | Współczynnik kontyngencji | ,45 | ,000 |

Cramer's V

Contingency Coefficient

W powyższej tabeli współczynnik zbieżności Pearsona wynosi C=0,45, a współczynnik zbieżności Cramera – V=0,51. Poziom istotności dla obydwu współczynników przekracza 0,000.

Crosstabs

Powyżej przedstawiono wszystkie tabele zrobione przez program Tabele krzyżowe. Zaprezentowane zostaną jeszcze dwa przykłady opracowania Tabele krzyżowe. W pierwszym przykładzie centralnym punktem jest informacja dotycząca oczekiwanych liczebności. Drugi przykład zawiera najmniejszą możliwą tabelę 2x2 (2 wiersze i 2 kolumny).

Poniższa tabela zawiera dane dla zmiennych z większą liczbą kategorii.

Tabela 27. Tabela dwóch zmiennych

tabela zawiera dwie zmienne

| | ankieta2 | | | | | |
|---|---|---|---|---|---|---|
| V5 | 2 | 3 | 4 | 5 | 6 | Ogółem |
| 2 | 5 | 38 | 55 | 37 | 2 | 137 |
| | 3,6% | 27,7% | 40,1% | 27,0% | 1,5% | 100,0% |
| 3 | 6 | 109 | 213 | 197 | 14 | 539 |
| | 1,1% | 20,2% | 39,5% | 36,5% | 2,6% | 100,0% |
| 4 | 3 | 67 | 284 | 311 | 33 | 698 |
| | ,4% | 9,6% | 40,7% | 44,6% | 4,7% | 100,0% |
| 5 | 0 | 17 | 77 | 181 | 41 | 316 |
| | ,0% | 5,4% | 24,4% | 57,3% | 13,0% | 100,0% |
| 6 | 0 | 1 | 4 | 31 | 9 | 45 |
| | ,0% | 2,2% | 8,9% | 68,9% | 20,0% | 100,0% |
| Ogółem | 14 | 232 | 633 | 757 | 99 | 1735 |
| | ,8% | 13,4% | 36,5% | 43,6% | 5,7% | 100,0% |

expected counts less than 5

W tak wielkich tabelach zdarza się często, iż oczekiwane liczebności są za niskie. Wartość chi-kwadrat wyliczona dla tej tabeli wynosi:

Tabela 28. Wyniki testu chi-kwadrat

Pearson Chi-Square

| | Wartość | df | Istotność asymptotyczna (dwustronna) |
|---|---|---|---|
| Chi-kwadrat Pearsona | 207,692 | 16 | ,000 |
| Iloraz wiarygodności | 199,500 | 16 | ,000 |
| Test związku liniowego | 172,712 | 1 | ,000 |
| N Ważnych obserwacji | 1735 | | |
| 20,0% komórek (5) ma liczebność oczekiwaną mniejszą niż 5. Minimalna liczebność oczekiwana wynosi 0,36. | | | |



Wartość chi-kwadrat wynosi 207,692 i ma poziom istotności 0,000. Jednak pod tabelą nr 28 znajduje się informacja, że 20% liczebności

minumum expected count

oczekiwanych jest niższych od 5 (oznacza to, że stosowanie testu jest jeszcze możliwe) i że najniższa liczebność oczekiwana wynosi 0,36 (to już nie pozwala na stosowanie testu).

W takich przypadkach należy znaleźć najniższą liczebność (uwaga: w komórkach tabeli znajdują się liczebności otrzymane, a nie oczekiwane!). Najniższa liczebność oczekiwana znajduje się w komórce tabeli na skrzyżowaniu kolumny z najniższą sumą „ogółem" i wiersza z najniższą sumą „ogółem" (czyli na dnie kolumny i na prawym krańcu wiersza). W tabeli jest to komórka na skrzyżowaniu kategorii 6 zmiennej wykształcenie (V5) i kategorii 2 w pytaniu ankietowym (czyli sumy: w prawo 45 i na dole 14). Oczekiwaną liczebność wylicza się w następujący sposób: mnoży się wymienione sumy „ogółem" i dzieli przez liczebność całej próby (w naszym przypadku: 45 razy 14 podzielić przez 1735=0,36).

łączenie kategorii

Aby pozbyć się liczebności 0,36 należy połączyć kategorie 5 i 6 w zmiennej wykształcenie. Wszystkie oczekiwane liczebności połączonej kategorii (w wierszu) zwiększą się. W przypadkach, gdy kategorie jednej zmiennej zbyt różnią się od siebie, co powoduje, że ich łączenie nie jest zasadne, można podjąć próbę połączenia kategorii w drugiej zmiennej. W omawianym przypadku są to kategorie 2 i 3 w pytaniu ankietowym. Jeżeli to połączenie też nie jest możliwe, należy wówczas wybrać inne rozwiązanie lub zrezygnować z zastosowania testu. Połączenie kategorii zmiennych porządkowych na ogół jest wykonalne, natomiast łączenie kategorii zmiennych nominalnych tylko incydentalne. Można łączyć też kilka kategorii w nową kategorię. Jeżeli zachodzi konieczność, łączenia można dokonać według obydwu zmiennych. Obowiązuje jednak zasada, że dokonuje się tylko tylu połączeń, które umożliwiają już przeprowadzenie testu.

rezygnacja z zastosowania testu

nowa tabela
nowe liczebności
nowe procenty
nowa ilość stopni wolności

Po połączeniu kategorii otrzymuje się nową tabelę – zmieniają się liczebności i procenty połączonych kategorii. Wszystkie wyliczone wskaźniki też będą inne: wartość chi-kwadrat, ilość stopni wolności (g lub df), poziom istotności, wartości współczynników zbieżności, itd.

Expected Count

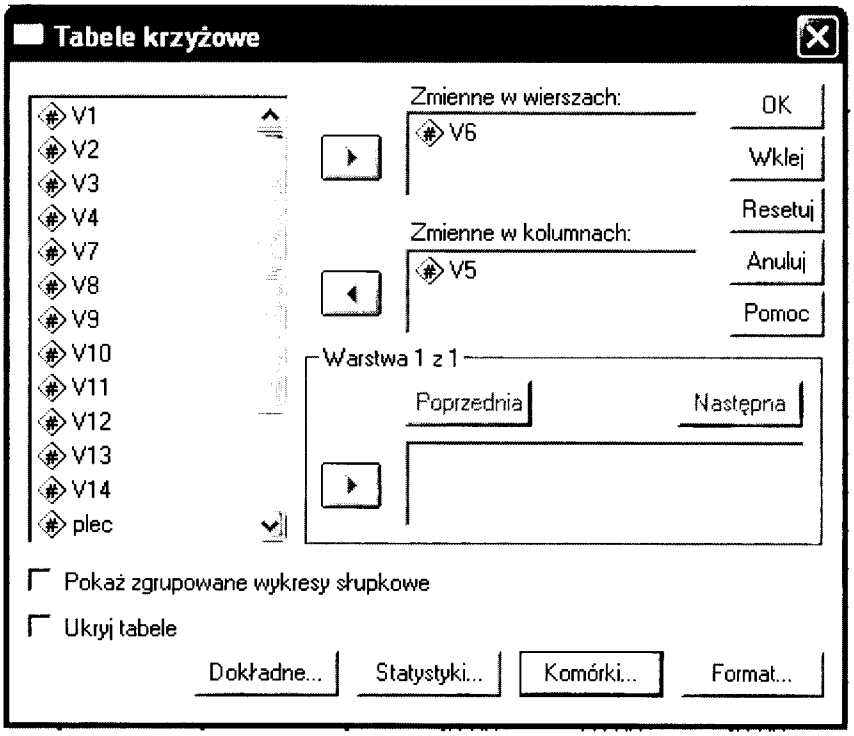
Significance

Odmienna sytuacja występuje w przypadkach, gdy już w pierwotnej tabeli wartość chi kwadrat nie jest większa od krytycznej. W takich przypadkach nie odrzuca się hipotezy zerowej i nie generalizuje się na populację generalną. Wówczas nie zachodzi konieczność przestrzegania reguł dotyczących oczekiwanych liczebności. W takich sytuacjach interpretuje się procenty z tabeli pierwotnej (wcale nie łączy się kategorii). Interpretacja ta odnosi się tylko do próby, a nie do całej populacji generalnej. Łączenie kategorii przeprowadza się więc tylko w przypadkach, gdy oczekiwane liczebności są za niskie i jednocześnie poziom istotności wynosi przynajmniej 0,05.

Crosstabs +Cells

Jeżeli próba badawcza nie jest liczna i zmienne posiadają większą liczbę kategorii (co pozwala przewidzieć, że liczebności oczekiwane będą za niskie), można już przy pierwszym użyciu podprogramu Tabele krzyżowe korzystać z opcji wyliczania oczekiwanych liczebności. Należy wówczas w oknie Tabele krzyżowe kliknąć na przycisk **Komórki**.

okno Crosstabs




Rysunek 46. Okno Tabele krzyżowe

Otworzy się nowe okno Tabele krzyżowe: Zawartość komórek.


Okno Crosstabs:Cells

Rysunek 47. Okno Tabele krzyżowe: Zawartość komórek

Kliknij na opcję **Oczekiwane** w ramce Liczebności. W tabeli rezultatów we wszystkich komórkach zostaną wyliczone oczekiwane liczebności. Wówczas jest łatwiej zdecydować, które kategorie należy połączyć. W końcowej tabeli należy wyłączyć opcje oczekiwanych liczebności, ponieważ do interpretacji wcale nie są one potrzebne (nawet odwrotnie: przeszkadzają!). Tabela zawierająca oczekiwane liczebności znajduje się na kolejnej stronie.

Expected Counts

tabela zawiera dwie zmienne

kilka kategorii

Tabela 29. Liczebności oczekiwane

| | | V5 | | | | | Ogółem |
|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | |
| V6 | 2 | 11,4 | 44,7 | 57,9 | 26,2 | 3,7 | 144,0 |
| | 3 | 32,2 | 126,8 | 164,1 | 74,3 | 10,6 | 408,0 |
| | 4 | 45,6 | 179,6 | 232,5 | 105,3 | 15,0 | 578,0 |
| | 5 | 41,1 | 161,5 | 209,2 | 94,7 | 13,5 | 520,0 |
| | 6 | 6,7 | 26,4 | 34,2 | 15,5 | 2,2 | 85,0 |
| Ogółem | | 137,0 | 539,0 | 698,0 | 316,0 | 45,0 | 1735,0 |

Podrozdział ten zamknie ostatni przykład, w którym wybrano tabelę dwóch zmiennych z dwoma kategoriami.

Tabela 30. Przykład najmniejszej możliwej tabeli

tabela zawiera dwie zmienne

dwa razy dwie kategorie

| | | udział ucznia w kółku szkolnym | | Ogółem |
|---|---|---|---|---|
| | | NIE | TAK | |
| płeć | dziewczynki | 390 | 506 | 896 |
| | | 43,5% | 56,5% | 100,0% |
| | chłopcy | 477 | 362 | 839 |
| | | 56,9% | 43,1% | 100,0% |
| Ogółem | | 867 | 868 | 1735 |
| | | 50,0% | 50,0% | 100,0% |

Zmienną niezależną jest płeć z kategoriami dziewczynki i chłopcy, a zmienną zależną jest udział ucznia w pewnym kółku szkolnym z kategoriami NIE i TAK. Jest to najmniejsza możliwa tabela.

Tabela 31. Wyniki testu chi-kwadrat

Chi-Square Test

Valid Cases

| | Wartość | df | Istotność asymptotyczna (dwustronna) |
|---|---|---|---|
| Chi-kwadrat Pearsona | 30,780 | 1 | ,000 |
| Poprawka na ciągłość | 30,249 | 1 | ,000 |
| N Ważnych obserwacji | 1735 | | |
| 0,0% komórek (0) ma liczebność oczekiwaną mniejszą niż 5. Minimalna liczebność oczekiwana wynosi 419,26. | | | |

Z tabeli wynika, że chi-kwadrat wynosi 30,780, a poziom istotności 0,000. Pod tabelą znajduje się informacja, że żadna z oczekiwanych liczebności nie jest niższa od 5. Jeżeli liczebności są za niskie, należy przeprowadzić test na podstawie skorygowanej wartości chi-kwadrat (30,249). Wartość ta jest otrzymana według korektury Yatesa (patrz Winner 1971).

Tabela 32. Współczynniki zbieżności

Phi-Coefficient

Contingency Coefficient

| | | Wartość | Istotność przybliżona |
|---|---|---|---|
| Nominalna przez Nominalna | Phi | ,133 | ,000 |
| | Współczynnik kontyngencji | ,132 | ,000 |
| N Ważnych obserwacji | | 1735 | |

W przypadkach tabel 2x2 można stosować współczynnik zbieżności Phi (0,133) lub zwyczajny współczynnik zbieżności Pearsona (C=0,132).

# Program Korelacje

Correlate

## *Podprogram Parami*

Bivariate

Na końcu tego podprogramu otrzymuje się współczynniki korelacji Pearsona i ich poziom istotności (w przypadkach prób). Rezultaty podprogramu Parami są prezentowane tak, jak w kolejnej tabeli.

Tabela 33. Współczynniki korelacji

| | | test1 | test2 | test3 |
|---|---|---|---|---|
| Pearson Correlations | test1 | 1 | ,728** | ,650** |
| | | | ,000 | ,000 |
| | | 1735 | 1735 | 1735 |
| | test2 | ,728** | 1 | ,452** |
| | | ,000 | | ,000 |
| | | 1735 | 1735 | 1735 |
| | test3 | ,650** | ,452** | 1 |
| | | ,000 | ,000 | |
| | | 1735 | 1735 | 1735 |

Wyliczono współczynnik korelacji dla każdej pary zmiennych, bez względu na to, które są niezależne, a które zależne. Tabela zawiera (zupełnie niepotrzebnie!) nawet współczynniki korelacji każdej zmiennej samej z sobą. Ich wartość oczywiście równa się r = 1,00. Z tabeli wynika, że korelacja np. między osiągnięciami w testach test1 i test2 wynosi r=0,728. Każda komórka zawiera też poziom istotności współczynnika korelacji i liczebność próby. W celu łatwiejszego przeglądu bardziej obszernych tabel, a następnie poszukiwania istotnych współczynników, dodano oznaczenia w postaci jednej lub dwóch gwiazdek. Jedną gwiazdką oznaczono wszystkie współczynniki, których poziom istotności wynosi 0,05, natomiast dwiema gwiazdkami zostały oznaczone współczynniki o poziomie istotności 0,01.

Sigtnificance

Partial

## *Opracowanie Cząstkowe*

W podobnie prosty sposób są prezentowane rezultaty korelacji cząstkowej (opracowanie Korelacje Cząstkowe). Przykład prezentuje tabela 34.

PartialCorrelations

Tabela 34. Współczynniki korelacji cząstkowej

| Zmienne kontrolne | | | test1 | test2 |
|---|---|---|---|---|
| test3 | test1 | Korelacja | 1,000 | ,641 |
| | | Istotność | . | ,000 |
| | | df | 0 | 1732 |
| | test2 | Korelacja | ,641 | 1,000 |
| | | Istotność | ,000 | . |
| | | df | 1732 | 0 |

Partial Correlation Coefficients

Przykład dowodzi, że cząstkowa korelacja między zmiennymi test1 i test2 wynosi r= 0,641 (gdy wyłączono wpływ zmiennej test3). Poziom istotności współczynnika korelacji cząstkowej wynosi 0,000 (jeszcze raz należy podkreślić, iż tak wysoki poziom to przede wszystkim rezultat wielkiej liczebności próby). Poprawny i konsekwentny zapis tego współczynnika korelacji wynosiłby wówczas:

Controlling for...

Significance

$r_{test1test2.test3}=0,641$

lub krócej:

$r_{12.3}=0,641$

W następnym przykładzie został obliczony współczynnik korelacji cząstkowej – tym razem trzeciego rzędu.

Tabela 35. Współczynniki korelacji cząstkowej trzeciego rzędu

| Zmienne kontrolne | | | test1 | test2 |
|---|---|---|---|---|
| test3 & test4 & test5 | test1 | Korelacja | 1,000 | ,626 |
| | | Istotność | . | ,000 |
| | | df | 0 | 1730 |
| | test2 | Korelacja | ,626 | 1,000 |
| | | Istotność | ,000 | . |
| | | df | 1730 | 0 |

Partial Correlation Coefficients

Współczynnik korelacji cząstkowej trzeciego rzędu między zmiennymi test1 i test2 przy wyłączeniu wpływu trzech zmiennych: test3, test4 i test5, wynosi 0,626. Poziom istotności współczynnika korelacji cząstkowej wynosi 0,000. Poprawny i konsekwentny zapis tego współczynnika korelacji wynosiłby wówczas:

$r_{test1test2.test3test4test5} = 0,626$

lub krócej:

$r_{12.345}=0,626$

Chi-Square Test

# Test zgodności

Rezultaty testu zgodności prezentowane są w dwóch tabelach. Pierwsza tabela zawiera liczebności, czyli dane początkowe. Wydawać by się mogło, że jest to najłatwiejsza część zadania. Jednak dużo więcej pracy wymaga tworzenie tabel niż wyliczenie wartości chi-kwadrat.

Observed N
Expected N
Residual

Pierwsza tabela zawiera liczebności otrzymane (otrzymane w ankiecie, wywiadzie itd.), liczebności oczekiwane i różnice tych liczebności (reszty). Przykład takiej tabeli dla zmiennej ankieta1 przedstawiono w tabeli 36.

Tabela 36. Liczebności w teście zgodności

| | Obserwowane N | Oczekiwane N | Reszty |
|---|---|---|---|
| 1 | 5 | 289,2 | -284,2 |
| 2 | 165 | 289,2 | -124,2 |
| 3 | 640 | 289,2 | 350,8 |
| 4 | 701 | 289,2 | 411,8 |
| 5 | 216 | 289,2 | -73,2 |
| 6 | 8 | 289,2 | -281,2 |
| Ogółem | 1735 | | |

Druga tabela zawiera wyniki testu zgodności: wartość chi-kwadrat (wartość wyliczoną lub empiryczną), liczbę stopni wolności i poziom istotności. Pod tabelą znajduje się informacja dotycząca liczebności oczekiwanych.

Tabela 37. Wyniki testu zgodności

| | Wartość |
|---|---|
| Chi-kwadrat | 1636,658 |
| df | 5 |
| Istotność asymptotyczna | ,000 |
| 0 komórek (0,0%) ma liczebność oczekiwaną mniejszą od 5. Minimalna liczebność oczekiwana w komórce wynosi 289,2 | |

Chi-Square
degrees of fredom
Significance
Expected Counts

Wartość chi-kwadrat wynosi 1636,658 a przy czterech stopniach wolności poziom istotności wynosi 0,000. Żadna z oczekiwanych liczebności nie jest mniejsza od 5.

Liczebności absolutne (liczba respondentów w okienku) nie są specjalnie przydatne do interpretacji, bowiem nie umożliwiają prostych i rzetelnych porównań. Sposób prezentacji rezultatów nie jest więc przemyślany. W pierwszej tabeli brak procentów do interpretacji. Zupełnie niepotrzebne do interpretacji są liczebności oczekiwane i różnice pomiędzy liczebnościami oczekiwanymi i otrzymanymi. Niestety, podprogram ten nie ma opcji obliczania procentów. Najlepszym rozwiązaniem będzie wykonanie zwyczajnej tabeli za pomocą podprogramu Częstości przed stosowaniem testu zgodności. Wówczas z dwóch tabel z rezultatami testu zgodności zostaje tylko druga tabela (Testy Chi-kwadrat).

Frequencies
Chi-Square Test

W tym opracowaniu należy omówić jeszcze jeden szczegół. Podobnie jak w teście zależności – pod główną tabelą Testy Chi-kwadrat pojawia się informacja o oczekiwanych liczebnościach. Informacja ta pojawia się przy obydwu testach, ale dostosowana jest tylko do pierwszego z nich. W teście zgodności wszystkie oczekiwane liczebności są identyczne. Jeżeli jedna liczebność jest za niska, to za niskie są jednocześnie wszystkie! Wystarczyłaby więc informacja, że oczekiwane liczebności są wyższe od 5. Procent za niskich liczebności jest już zupełnie niepotrzebny, bo i tak może wynosić jedynie 0% (tak, jak w naszym przypadku) lub 100%.

# Zapisywanie i ponowne korzystanie z plików powstałych w SPSS

*Z tego rozdziału nauczysz się:*

- *zapisywać pliki z uporządkowanymi danymi*
- *zapisywać pliki z poleceniami dla opracowań*
- *zapisywać pliki z wynikami*
- *odtwarzać te pliki i ponownie z nich korzystać*
- *eksportować wyniki*

# Wstęp

Zdarza się, że opracowanie w SPSS jest proste i krótkie. Wówczas dokonuje się wszystkiego już za pierwszym razem i nie jest koniecznie zapisywanie plików. W prostszych przypadkach nie jest potrzebna wiedza o sposobach zapisywania i ponownego korzystania z SPSS plików.

Zapisywanie jest koniecznością, jeżeli wszystko nie zostało zrobione przy pierwszym uruchomieniu SPSS. W takich przypadkach należy zapisać i zachować wszystkie pliki. W niniejszym rozdziale zostaną opisane procedury zapisywania i ponownego uruchamiania plików. W programie SPSS powstają trzy rodzaje plików:

Data File
- pliki danych (rozszerzenie sav),

Syntax File
- pliki poleceń (rozszerzenie sps),

Output File
- pliki rezultatów (rozszerzenie spo).

Najbardziej potrzebne i przydatne są pliki danych. Pliki poleceń są zazwyczaj najmniej potrzebne, jednak zostaną omówione, ponieważ zabierają wyjątkowo mało miejsca na dyskach.

Pliki rezultatów są potrzebne tymczasowo – dopóki wszystkie rezultaty nie zostaną wykorzystane.

Dyski w nowoczesnych komputerach mają tyle miejsca, że kilkadziesiąt plików powstałych w SPSS (a nawet i kilka tysięcy!) nie przedstawia zauważalnego obciążenia. Dlatego też należy korzystać z tych możliwości i bez zastanawiania się zapisywać wszystkie pliki. Już na zwyczajnej dyskietce można zapisać dużo plików. Zostanie to zilustrowane przykładem stosunkowo dużego opracowania:

- W pliku danych znajdują się dane i opisy 12 zmiennych dla 1735 respondentów – zapisany plik ma objętość zaledwie 20 KB.

Frequencies
- W pliku poleceń zapisano opracowanie Częstości dla wszystkich zmiennych (dodatkowo wybrano obliczanie wszystkich wskaźników) - zapisany plik ma objętość tylko 1 KB.

- W pliku rezultatów znajdują się wszystkie wyniki omówionego opracowania: tabele i wyliczone wskaźniki – zapisany plik ma objętość 69 KB.

objętość mniejsza niż 100 KB

Wszystkie trzy pliki razem mają objętość mniejszą niż 100 KB – więc na dyskietce można zapisać kilkanaście takich opracowań. Należy jednak nadmienić, że plik z rysunkami byłby znacznie większy. Plik z rezultatami jest potrzebny tylko tymczasowo. Po przeniesieniu wyników do tekstu badania (np. do raportu badania, artykułu, książki lub pracy magisterskiej itd.) plik można wykasować. To samo opracowanie można wykonać kiedykolwiek ponownie, ponieważ plik poleceń został zapisany. Zachowując tylko niezbędne pliki, na jednej dyskietce można umieścić pliki z prawie stu podobnych opracowań. Jedna dyskietka mogłaby wystarczyć na badania mniejszego instytutu, wykonane w ciągu całego semestru!

dyskietka wystarczy na badania całego instytutu

## Przygotowanie do zapisu

Z nazwami plików w SPSS jest trochę więcej pracy niż w innych programach, ponieważ powstają trzy rodzaje plików. Aby zapobiec chaosowi, należy podjąć następujące działania:

nowy folder

- Dla każdego badania zrobić nowy folder jeszcze przed rozpoczęciem pracy z SPSS. W tym folderze umieszcza się wszystkie pliki dotyczące tego badania.

Data File

- Pierwszym plikiem będzie plik danych powstały w programie Word. Ten plik posiada rozszerzenie txt. Będzie to np. plik: ankieta.txt. Jeżeli w pewnym badaniu zastosowano kilka ankiet (które w dodatku opracowuje się odrębnie), najlepszym rozwiązaniem jest założenie dla każdej ankiety nowego folderu. Podobnie też w przypadku, gdy w badaniu pojawia się kilka prób (np. nauczyciele, uczniowie i rodzice).
- Wszystkie pliki powstałe w trakcie pracy z SPSS zapisuje się w tym folderze.

- Do nazw plików rozsądnie jest dodawać kolejne cyfry (uwaga: w nazwie, a nie w rozszerzeniu).

Data Files

- Wszystkim plikom danych powstałych w SPSS w tym folderze najlepiej nadać taką samą nazwę, ale z dodatkową cyfrą. Powstanie w ten sposób szereg plików np.: uczniowie1.sav, uczniowie2.sav, uczniowie3.sav itd. lub dane1.sav, dane2.sav, itd.

Syntax Files

- Wszystkim plikom poleceń powstałych w SPSS w tym folderze najlepiej nadać taką samą nazwę, ale z dodatkową cyfrą (lub datą). Powstanie w ten sposób szereg plików np.: Polecenia1.sps, Polecenia2.sps, Polecenia3.sps itd. lub Polecenia25lutego.sps, Polecenia9marca.sps, itd.

Output Files

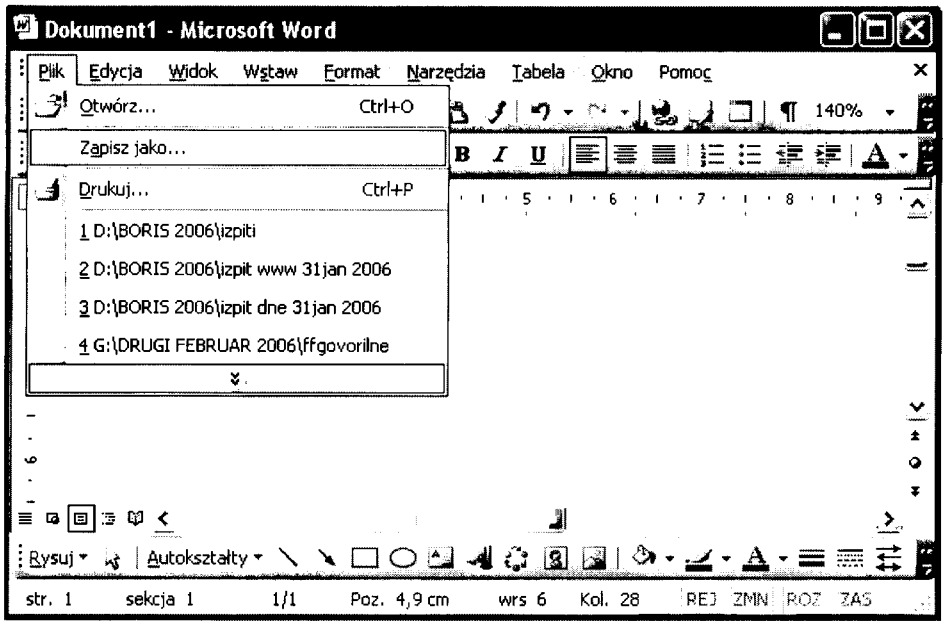
- Wszystkim plikom rezultatów w tym folderze najlepiej nadać taką samą nazwę, ale z dodatkową cyfrą (lub datą). Powstanie w ten sposób szereg plików np.: Raport1.spo, Raport2.spo, Raport3.spo itd. lub Raport25lutego.spo, Raport9marca.spo, itd.

# Zapis pliku danych

Jeden plik danych znajduje się już w tym folderze. Po wniesieniu danych przy pomocy programu Word, plik został zapisany (w TXT formacie). Zawiera on jednak tylko surowe dane. W trakcie importu do SPSS dane te zostają szczegółowo opisane (np. imię, rodzaj i położenie zmiennych, brakujące dane itd.). Jeżeli po ukończeniu importu plik nie został zapisany przy pomocy SPSS, przy każdym kolejnym korzystaniu z SPSS należałoby ponownie dokonać pełnego importu. Dlatego też najlepiej plik danych zapisać i zachować w folderze przed pierwszym opracowaniem. Powstanie w ten sposób nowy plik danych, który różni się od pliku powstałego w programie Word. Oto opis procedury zapisu.

File Save lub Save As

Po zakończeniu importu pojawi się sieć danych (rysunek 16). Aby zapisać plik należy otworzyć menu **Plik** i kliknąć na poleceniu **Zapisz** lub **Zapisz jako.**

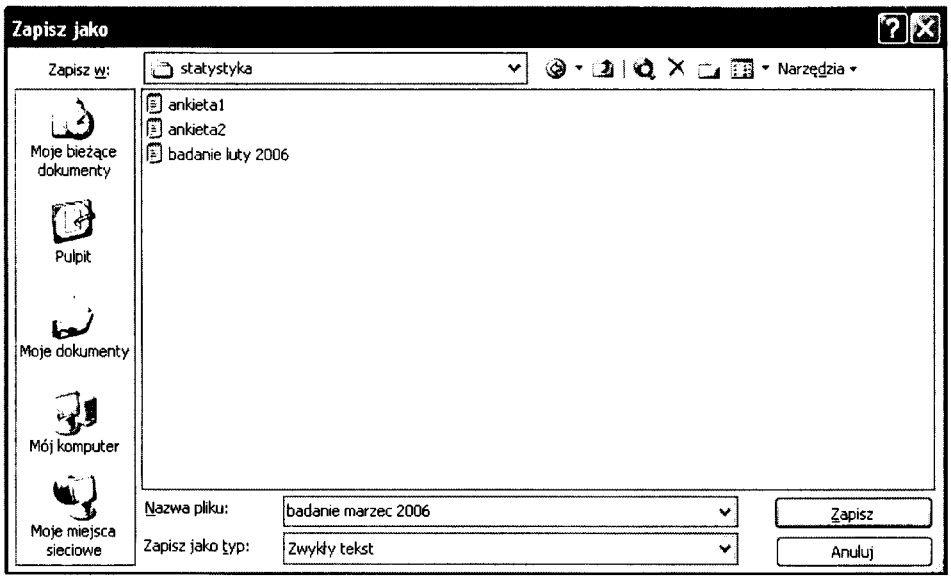



File

Save As...

Rysunek 48. Polecenie Zapisz jako

Drugie polecenie jest lepsze. Pojawi się okno:




Save in:

Okno Save Data As...

File name:
Save as type:

Rysunek 49. Okno Zapisz jako

Save as type: W dolnym polu (Zapisz jako typ) pojawi się automatyczne ustawienie: SPSS (z rozszerzeniem sav). Okno Zapisz jako jest nam już znane i nie będziemy szczegółowo opisywać procedury. Należy nadać nazwę pliku (np. dane1), natomiast wpisywanie rozszerzenia nie jest konieczne - Save program sam go doda. Na końcu należy przycisnąć przycisk **Zapisz.** Plik zostanie zapisany na dysku i można rozpocząć opracowywanie danych.

> ***Uwaga!***
>
> Transform Data Jeżeli w trakcie uruchamiania różnych opracowań dane zostały dodatkowo uporządkowane (np. przy pomocy poleceń w menu Przekształcenia), należy przed zamknięciem programu SPSS plik danych zapisać ponownie (bo nie jest to już ten sam plik).

Należy sobie uświadomić także:

- Save Jeżeli stara (niezmieniona) wersja pliku nie jest nam potrzebna, zapisu dokonuje się przy pomocy polecenia **Zapisz** (lub przy pomocy przycisku z ikoną dyskietki lub korzystając z klawiszy **Ctrl** i **S** na klawiaturze). W ten sposób zostanie zapisana nowa wersja, a stara będzie automatycznie usunięta. Jeżeli starej wersji naprawdę już nie potrzebujemy, to dobrze, że nie została ona zachowana, ponieważ mnóstwo podobnych plików stwarza jedynie chaos.
- Save As... Jeżeli stara wersja jest nam rzeczywiście potrzebna, należy skorzystać z polecenia **Zapisz jako** (ale w żadnym przypadku nie z przycisku z ikoną dyskietki lub z klawiszy **Ctrl** i **S**!!!). Należy kliknąć na **Zapisz jako**, w nazwie pliku wprowadzić dodatkową cyfrę, a wszystkie pozostałe ustawienia zostawić niezmienione. Save Na końcu należy przycisnąć **Zapisz**. Wówczas w folderze znajdą się obydwa pliki (np. uczniowie1.sav i uczniowie2.sav).

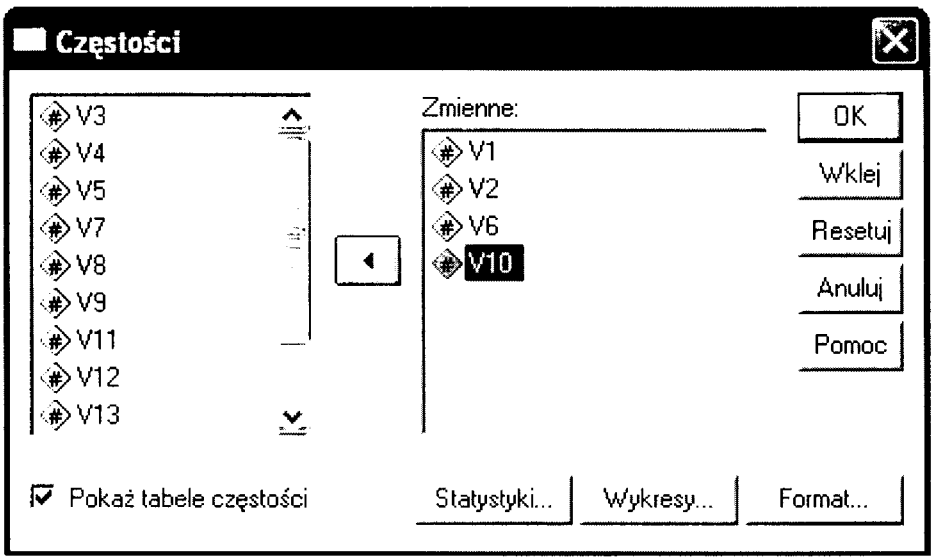
Jeżeli dane zamierza się opracować jeszcze w innym programie (Excell, Save As... DBase itd.), po zapisie pliku należy ponownie otworzyć okno **Zapisz** Save as type: **jako,** kliknąć w polu Zapisz jako typ na trójkąt i wybrać zapis w innym formacie.

# Zapis pliku poleceń

Syntax File

Po wyborze pewnego opracowania (i ustawieniu parametrów), startuje się przez **OK.** Przed naciśnięciem funkcji OK wszystkie polecenia można zapisać w pliku i zachować na dysku (zapis przebiegu opracowania). Na poniższym rysunku znajduje się np. okno do opracowania Częstości:

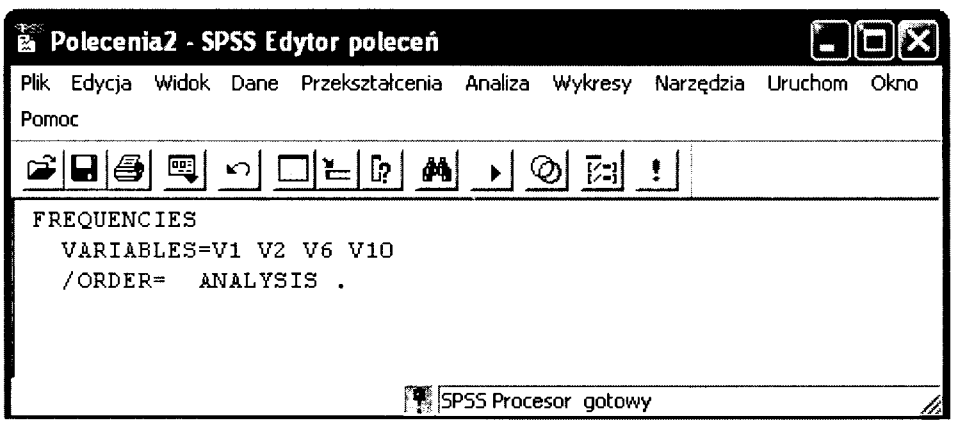
Frequencies




Rysunek 50. Okno Częstości

Przed naciśnięciem na **OK** należy przycisnąć na polecenie **Wklej.** Otworzy się okno SPSS Edytor poleceń:

Paste

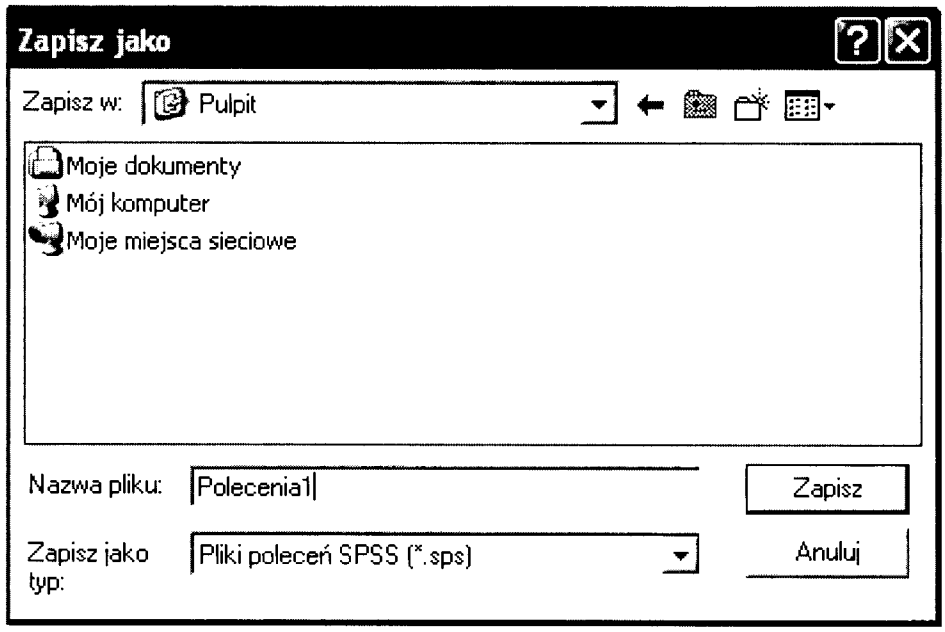



okno Syntax Editor

Rysunek 51. Okno Edytor poleceń

Save As...

W tym oknie został zapisany pełny przebieg opracowania (wszystkie polecenia w właściwej kolejności). Wybiera się opcje **Zapisz jako** i według znanej już procedury zachowuje się plik z tym zapisem:




Save as type

Rysunek 52. Okno Zapisz jako

W polu Zapisz jako typ został już ustawiony format Pliki poleceń SPSS. Automatycznie jest też ustawiona nazwa pliku Polecenia1. Program sam dodaje rozszerzenie sps.

Save

Na końcu należy potwierdzić wybrane ustawienia przyciskiem **Zapisz**. Plik zostanie zapisany. To samo opracowanie można uruchomić kiedykolwiek ponownie. Zamiast nazwy Polecenia można użyć nazwy opracowania, dodając na wszelki przypadek datę lub dodatkową cyfrę. W ten sposób powstają np. pliki: czestosci1.sps, korelacje1.sps, korelacje2.sps, krzyzowe23lutego.sps, krzyzowe5maja.sps itd.

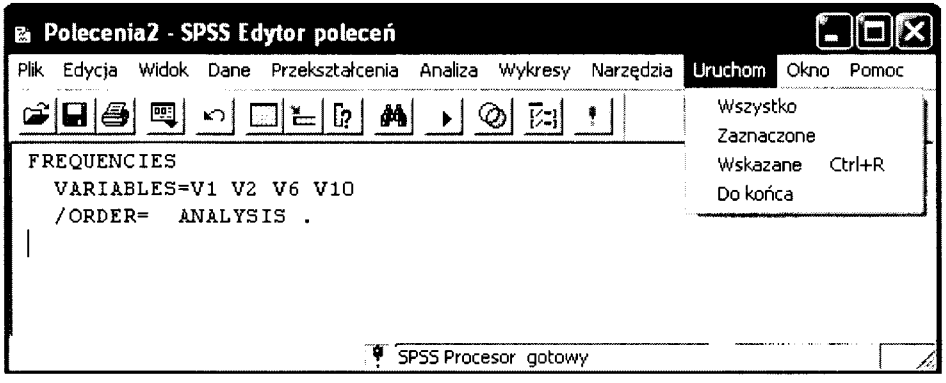
Frequencies

Należy jeszcze uruchomić zapisane opracowanie. Przed zapisywaniem i zachowaniem pliku poleceń było otwarte okno dialogowe przygotowanego opracowania (np. Częstości). W oknie tym należało

uruchomić opracowanie przy pomocy polecenia **OK.** Po zapisie opracowania zamknęło się to okno, a pojawiło okno Edytor poleceń. Aby uruchomić opracowanie należy otworzyć menu **Uruchom**, a następnie wybrać jedną z proponowanych opcji: Wszystko, Zaznaczone, Wskazane lub Do końca. W najprostszych przypadkach wystarczy opcja Wszystko.

OK

Run
All
Selected
Current
To End




Rysunek 53. Menu Uruchom

Aby zrozumieć proponowane opcje i właściwie z nich korzystać, należy zapoznać się z przebiegiem zapisu pliku poleceń w trakcie pracy nad SPSS. Każde nowe opracowanie przygotowane dla tych samych danych i zapisane przy pomocy opcji Wklej, dopisuje się na końcu tego samego pliku. Plik poleceń z każdym nowym opracowaniem staje się coraz dłuższy.

Plik poleceń z każdym nowym opracowaniem jest coraz dłuższy

Można to zilustrować przykładem: w oknie z danymi jako pierwsze zostało ustawione opracowanie Częstości i zostało zapisane przez opcje Wklej. Program SPSS otwiera nowe okno Edytor poleceń. W oknie tym zostały zapisane wszystkie polecenia oraz ustawienia opracowania Częstości przez opcje Zapisz jako. Aby uruchomić to opracowanie (bo właśnie po to zostało ono przygotowane) skorzystano z poleceń Uruchom i Wszystko. Opracowanie zostało wykonane i SPSS otworzył okno Edytor raportów. W przypadku, gdy po przeglądzie wyników chcemy wykonać jeszcze opracowanie Korelacje, należy wrócić do okna

Frequencies
Syntax Edytor

Frequencies
Save As...
Run+All
Output Viewer

z siecią danych i ustawić to nowe opracowanie. Aby je zapisać należy skorzystać z opisywanej opcji Wklej - znowu pojawi się okno Edytor poleceń. Do zapisu pierwszego opracowania dodano zapis drugiego. Korzystając wówczas z opcji Uruchom i Wszystko, program SPSS wykonuje ponownie pierwsze, a następnie drugie opracowanie. W pliku wyników dwa razy pojawiają się wyniki pierwszego opracowania. Jeżeli już wykonaliśmy trzy opracowania, to chaos w wynikach jest coraz większy. Z tego powodu dodano pozostałe opcje w menu Uruchom (Zaznaczone, Wskazane i Do końca). Można z nich wybrać:

Paste

Run+All

Run

All - **Wszystko** – w przypadku, gdy pracujemy nad pierwszym opracowaniem.

Selected - **Zaznaczone** – w przypadku, gdy należy uruchomić tylko wybraną część opracowania. W pliku poleceń należy oznaczyć wybraną część i kliknąć polecenie Zaznaczone. Uruchomi się wówczas tylko wybrana część.

Current - **Wskazane** – w przypadku, gdy ma być wykonane tylko opracowanie dopisane przy ostatnim użyciu opcji Wklej.

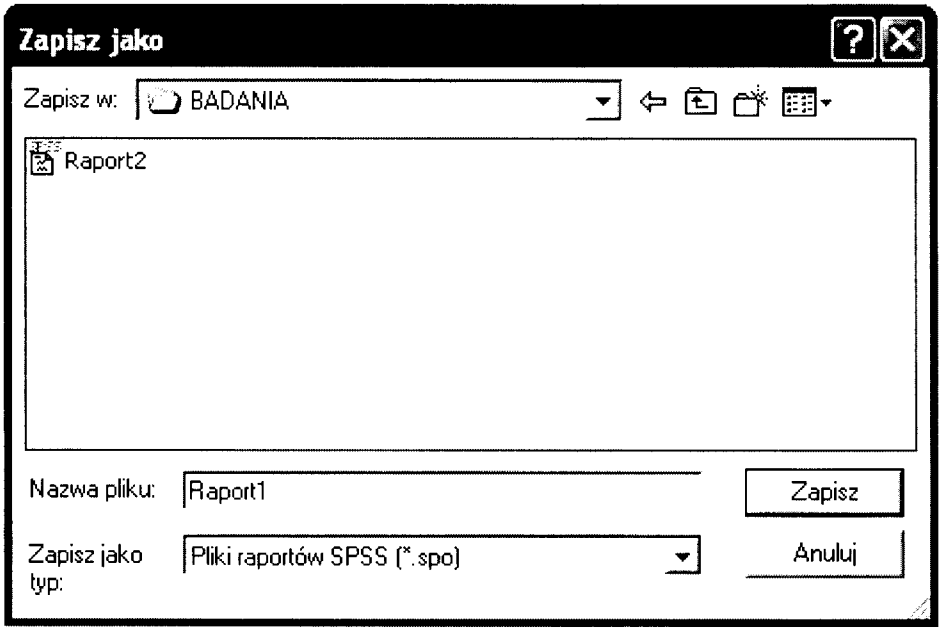
To End - **Do końca** – w przypadku, gdy ma być uruchomione wszystko od pewnego miejsca do końca pliku Polecenia. Początek jest tam gdzie jest kursor.

## Zapis pliku wyników

Jeżeli niektóre rezultaty zostały niewykorzystane, należy zachować plik z rezultatami na twardym dysku. Umożliwia to ponowne korzystanie z tych rezultatów w dowolnym czasie. Przechowywanie przez dłuższy czas jest jednak bezsensowne, ponieważ kiedykolwiek można otworzyć plik danych oraz plik poleceń i ponownie uruchomić te same opracowania. Wówczas po kilku sekundach otrzymuje się te same wyniki.

Output Viewer
File+Save As...

Rezultaty zachowuje się w oknie Edytor raportów. Należy otworzyć menu **Plik** i kliknąć na pozycji **Zapisz jako**. Pojawi się wówczas okno:




okno Save As

Rysunek 54. Okno Zapisz jako

W polu Zapisz jako typ należy zostawić ustawione automatycznie Pliki raportów (z rozszerzeniem spo). Automatycznie jest także ustawiona nazwa pliku Raport1. Najlepiej ustawienia te zostawić. Jeżeli powstanie więcej plików z rezultatami, można w nazwie dodać za każdym razem większą liczbę lub datę, albo też jedno i drugie. Data powstania pliku i tak jest zarejestrowana, tylko że nie jest widoczna we wszystkich programach.

Save as type:
Viewer Files (*.spo)

# Eksport pliku wyników

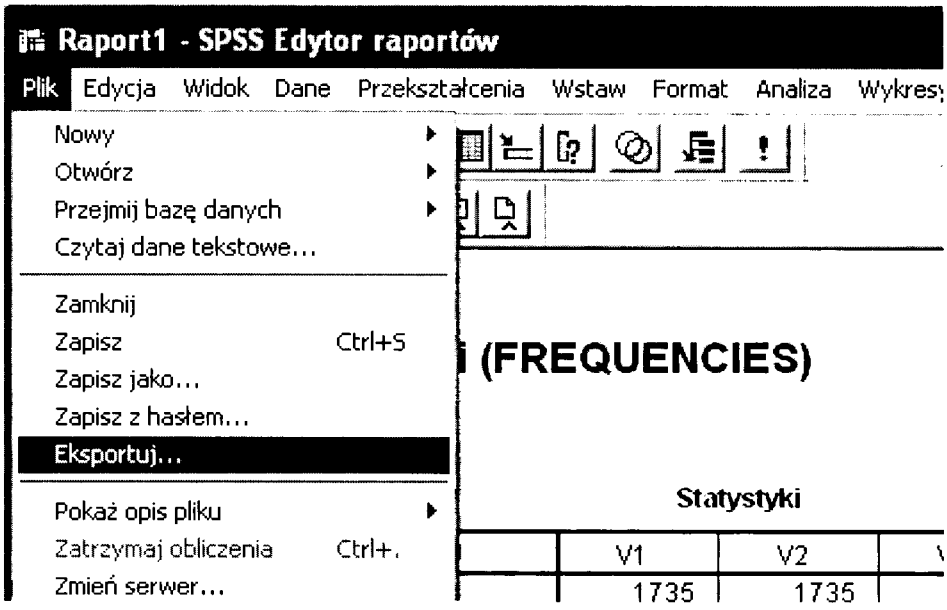
Eksport pliku jest potrzebny w przypadkach, gdy z wyników zamierza się korzystać w innych programach oprócz SPSS. Eksport jest konieczny, jeżeli końcowy tekst pisze się na komputerze, w którym nie ma SPSS (np. dane zostały opracowane w pracy, a tekst piszemy w domu). Rezultaty opracowania statystycznego należy eksportować i zapisać na dyskietce. W innym komputerze wyeksportowany plik wnosi się (Wstaw) do pliku w programie Word (np. do tekstu pracy magisterskiej).

Insert

Procedura zaczyna się w oknie Edytor raportów. Należy otworzyć menu **Plik** i wybrać polecenie **Eksportuj.**

File
Export...

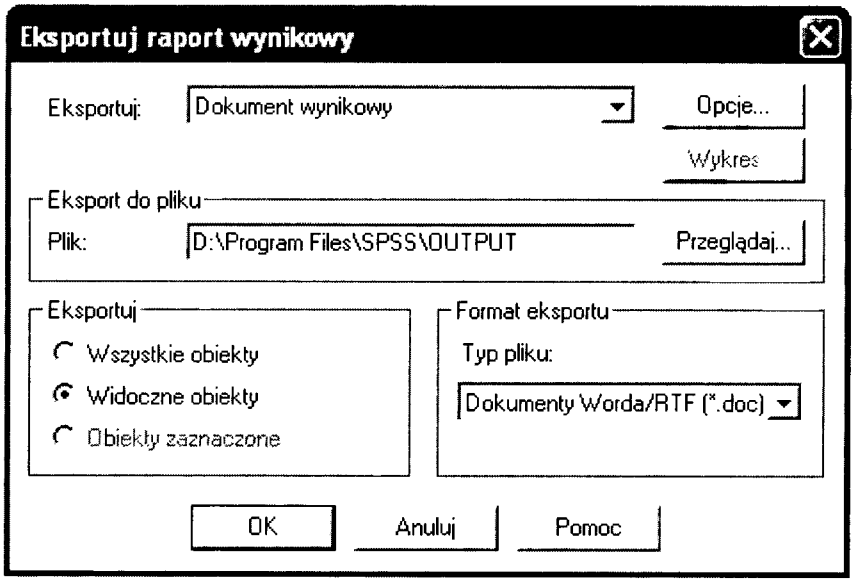
okno Output Viewer




Rysunek 55. Polecenie Eksportuj

Otworzy się okno do eksportu.

okno Export Output




Rysunek 56. Okno Eksportuj raport wynikowy

W pierwszym polu Eksportuj istnieją trzy opcje do wyboru.

Export:

1. Jeżeli rezultaty zawierają tabele i rysunki, należy wybrać opcję Dokument wynikowy.

Output Document

2. Jeżeli rezultaty zawierają tylko tabele, należy wybrać opcję Dokument wynikowy (bez wykresów).

Output Document (No Charts)

3. Jeżeli rezultaty zawierają tylko rysunki (grafikony), należy wybrać opcję Wyłącznie wykresy.

Charts Only

W drugim polu (Eksport do pliku) przy pomocy przycisku **Przeglądaj** należy określić miejsce, gdzie ma być zachowany plik z rezultatami (Przeglądaj jest opcją znaną użytkownikowi z innych programów w Windows). W trzecim polu Eksportuj zwykle wybiera się Widoczne obiekty lub Obiekty zaznaczone. Przy wyborze opcji Wszystkie obiekty, w zachowanym pliku znajdują się pewne tabele, które służą tylko jako informacja o przebiegu opracowania (nawet w Edytorze raportów są niewidoczne). Do interpretacji tabele te wcale nie są potrzebne.

Export File
File Name
Browse
Export what
All Objects
All Visible Objects
Selected Objects

W polu Format exportu należy określić, w jakiej postaci (formacie) plik ma być zachowany. Szczególnie przydatne są dwa formaty: Word dokument DOC i format HTML. Jeżeli dane opracowuje się po to, aby rezultaty umieścić w tekście (pracy magisterskiej, licenciackiej, dyplomowej, seminaryjnej itd.) najlepiej wybrać Word dokument DOC. Przydatny jest też format HTML, ponieważ przy przenoszeniu pliku do programu Word tabele będą poprawne i łatwe do korektur, a w dodatku plik w formacie HTML można umieścić bezpośrednio w internecie.

Export Format

File Type

Po ustawieniu wymienionych opcji należy kliknąć na **OK**, aby zachować plik. Z pliku HTML (rozszerzenie htm) można też przenosić poszczególne tabele do programu Word (przy pomocy poleceń Kopiuj i Wklej). Bezpośrednio z okna Edytor Raportów można też przenieść do programu Word poszczególne tabele. Należy oznaczyć tabelę w Edytorze Raportów i nacisnąć polecenie Kopiuj. Następnie trzeba otworzyć odpowiedni plik w programie Word i na określonym miejscu (np. w rozdziale Interpretacja wyników pracy magisterskiej) skorzystać z polecenia Wklej.

OK

Copy+Paste

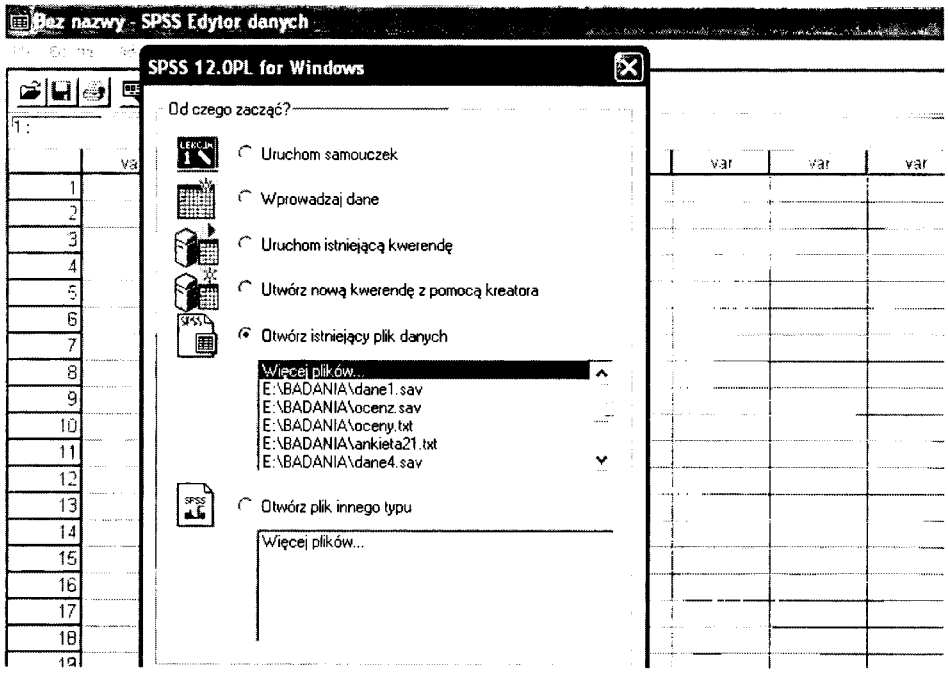
Data File

# Ponowne użycie pliku danych

Często nie udaje się wykonać wszystkich potrzebnych opracowań przy pierwszym podejściu. Należy wtedy ponownie otworzyć plik danych i kontynuować pracę.

Po uruchomieniu programu SPSS pojawi się okno z listą ostatnich dziewięciu zapisanych plików SPSS:

What would you like to do?



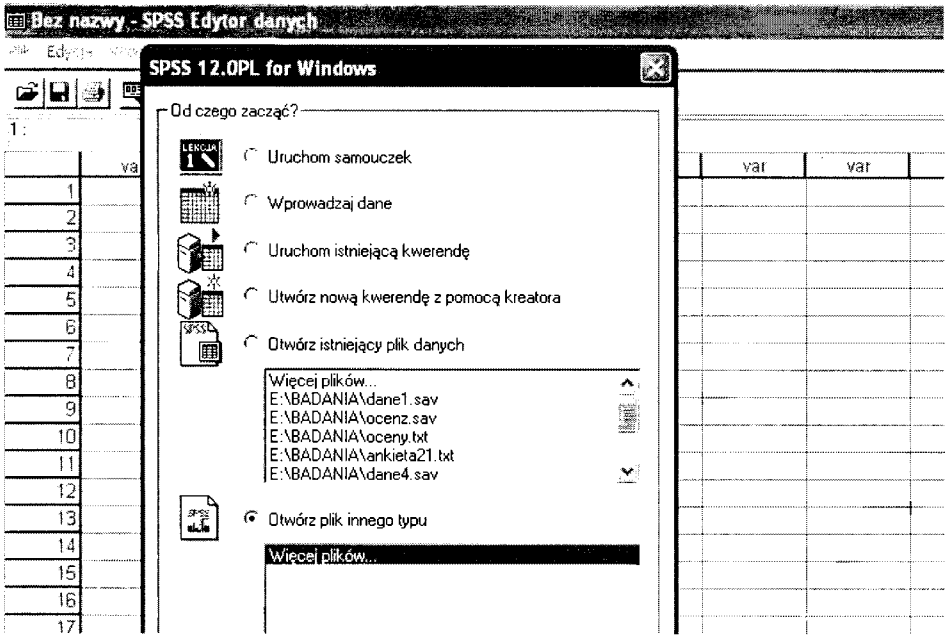

Rysunek 57. Okno powitalne programu SPSS

Open an existing file

Jeżeli na liście znajduje się plik danych, które należy dodatkowo opracować, wybiera się go i potwierdza kliknięciem na **OK**. Pojawi się sieć danych i można kontynuować prace nad nowymi opracowaniami.

More Files

Trudniej jest w przypadku, gdy plik z danymi nie znajduje się na liście proponowanych plików. W tym przypadku należy wybrać pierwszą kolejną propozycję poszukiwania pozostałych plików (Więcej plików...).

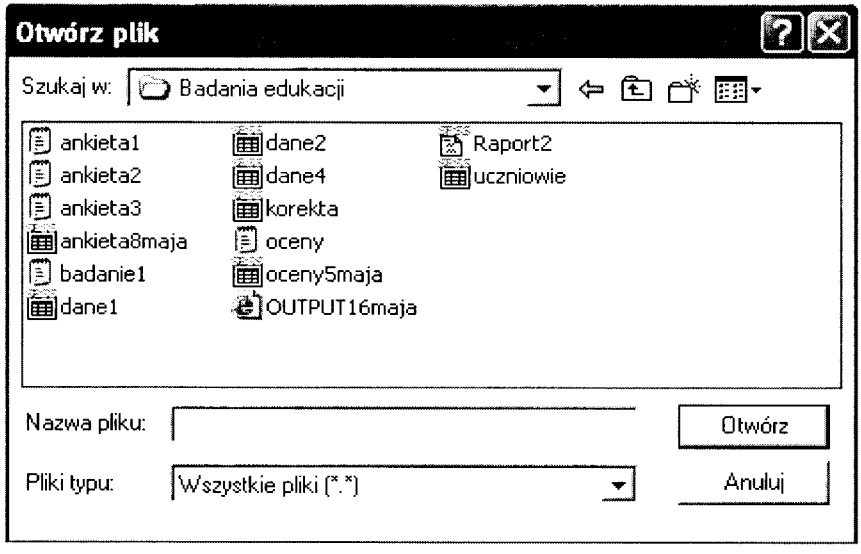



What would you like to do?

Rysunek 58. Okno powitalne programu SPSS

Należy wybrać tę opcję i kliknąć na **OK**. Pojawi się zwyczajne okno do poszukiwania plików.

More Files+OK




okno Open File

Rysunek 59. Okno Otwórz plik

Aby znaleźć ten plik, należy wiedzieć, gdzie się on znajduje i rozpoznać jego nazwę. Pliki danych posiadają w nazwie za kropką rozszerzenie sav. Niestety, okno to nie pokazuje pełnych nazw (brakuje akurat rozszerzeń!), lecz pliki są pokazane z ikoną i nazwą do kropki. Plik można rozpoznać tylko według ikony. Na rysunku w folderze Badania edukacji znajduje się siedem plików z danymi w SPSS formacie sav (ankieta8maja, dane1, dane2, dane4, korekta, oceny5maja i uczniowie). Należy wybrać odpowiedni plik i kliknąć na **Otwórz**.

Open

W tym folderze znajdują się też inne pliki: pięć plików z danymi tekstowymi (ankieta1, ankieta2, ankieta3, badanie1 i oceny), jeden plik z wynikami w SPSS formacie spo (Raport2) i jeden plik z wynikami w formacie HTML (Output16maja).

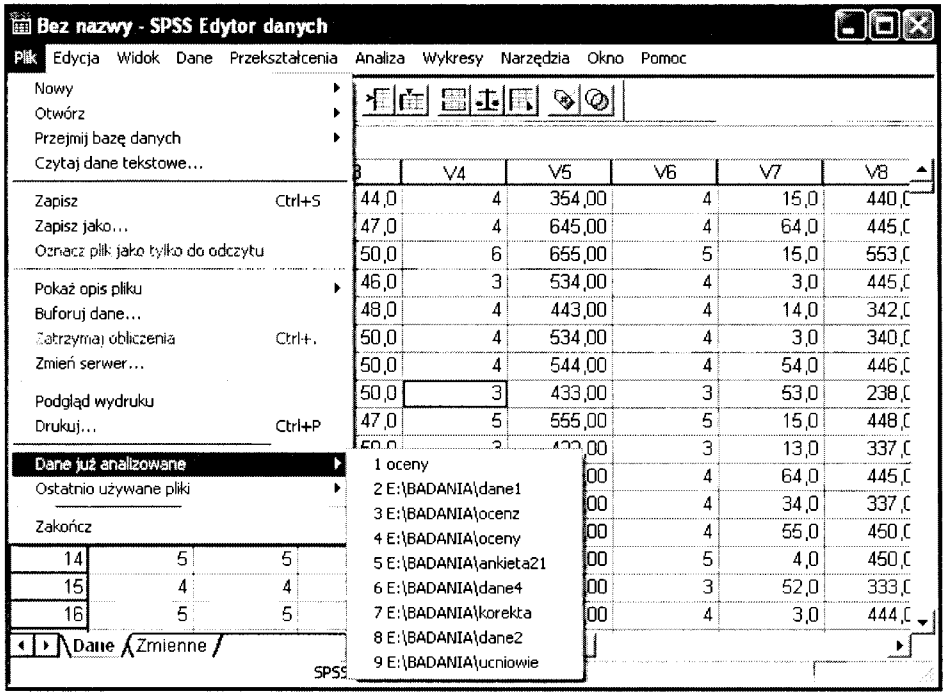
Trzeci sposób otwierania starych plików z danymi jest następujący: w podstawowym oknie SPSS otworzyć menu **Plik** i kliknąć na poleceniu **Otwórz**. Pojawi się wówczas znane już okno do poszukiwania plików. Rozszerzenie sav jest automatycznie ustawione. Należy poszukać odpowiedniego pliku i kliknąć dwa razy.

Menu File+ Open

> ***Uwaga!***
>
> Po kliknięciu na pozycji Otwórz w oknie z siecią danych (Edytor danych), program automatycznie ustawia rozszerzenie sav. Jeżeli wykonuje się to w oknie z wynikami (Edytor raportów), ustawi się rozszerzenie spo, a jeżeli w oknie poleceń (Syntax Editor), ustawi się rozszerzenie sps. Dlatego też najpierw należy ustawić rodzaj pliku na sav, a dopiero potem rozpocząć poszukiwania pliku.

Istnieje jeszcze jedna opcja, którą Czytelnik zna np. z programu Word: na dnie menu Plik znajduje się lista kilku ostatnich otwieranych plików. Niestety, na liście znajdują się różne pliki SPSS bez rozszerzeń i trudno rozpoznać odpowiedni plik. Z tej opcji należy korzystać w przypadkach, jeżeli precyzyjnie wiadomo, czego się poszukuje.

Menu File




okno Data Editor

Rysunek 60. Polecenie Dane już analizowane

***Uwaga!***

Jeżeli w trakcie opracowań zmieniono coś w danych (np. korzystając z Braki danych, Rekoduj, Oblicz wartości, itd.), należy przed końcem prac w SPSS ponownie zachować plik z danymi. W przeciwnym przypadku wszystkie nowe zmiany będą utracone. Jednak SPSS sam stara się o to, aby do tego nie doszło. Jeżeli w danych dokonano jakichkolwiek zmian, przed zamykaniem SPSS pojawi się pytanie, czy zmiany mają być zachowane. To jest ostatnia okazja, aby nie utracić nowych zmian, więc należy wówczas kliknąć **TAK**. Użytkownicy często odpowiadają na takie pytania powierzchownie lub nawet bez dokładnego czytania. Należy wierzyć programowi, że pytania te są naprawdę sensowne – program nie stosuje ich niepotrzebne! Jeżeli nie dokonano żadnych zmian lub, jeżeli zmiany już zostały zachowane, takie pytanie nie pojawi się.

Missing Values
Recode
Compute

Yes

Syntax

## Ponowne użycie pliku poleceń

Po otwarciu pliku danych można ponownie uruchomić opracowanie (pod warunkiem, że przebieg opracowania został zachowany w pliku Polecenie.sps).

Menu File+Open

W sieci danych (okno Edytor danych) należy otworzyć menu **Plik** i kliknąć na pozycji **Otwórz**. Otworzy się podmenu, w którym należy

Syntax

wybrać **Polecenia**. Pojawi się zwyczajne okno do poszukiwania plików. W rubryce Pliki typu ustawiono automatycznie Pliki poleceń SPSS. Następnie należy poszukać folderu swojego badania i w nim wybrać plik, w którym zostały zapisane wszystkie polecenia danego opracowania. Pojawi się okno Edytor poleceń. Po upewnieniu się, że wybrano właściwy plik, należy ponownie uruchomić to opracowanie

Run

przy pomocy polecenia **Uruchom**.

## Ponowne użycie pliku wyników

Zachowany wcześniej plik wyników (Raport.spo) można kiedykolwiek ponownie otworzyć.

Po uruchomieniu programu SPSS w oknie powitalnym należy znaleźć dany plik. Jeżeli plik ten nie znajduje się na proponowanej liście, należy

More Files+OK

wybrać polecenie **Więcej plików** i kliknąć na **OK**. Pojawi się zwyczajne okno do szukania plików. W rubryce Pliki typu należy wybrać Pliki

Files of type: Viewer Files

raportów, a następnie poszukać pliku w folderze badania. Po kliknięciu na nazwę pliku i po odczekaniu kilku sekund, otworzy się okno z wynikami.

# *Literatura*

## *Literatura*

Anderson R. B., *STAT POWER, An Apple computer program*, Cambridge 1981.

Borenstein M. i Cohen J., *Statistical power analysis: A computer program*, Hillsdale 1988.

Chraska M., *Empirická pedagogicka šetřeni a jejich statisticke vyhodnocovani*, Olomouc 1988.

Claus G. i Ebner H., *Podstawy statystyki dla psychologów, pedagogów i socjologów*, Warszawa 1972.

Cohen J., *Statistical power analysis for the behavioral sciences*, Hillsdale 1988.

Finch J., *Research and Policy, The uses of Qualitative Methods in Social and Educational Research*, London 1986.

Fox J. D., *The research process in Education*, New York 1969.

Góralski A., *Podstawowe metody statystyczne w psychologii i pedagogice*, Warszawa 1994.

Haase R. F., *A Basic program to compute atypical values of alpha*, Educational and Psychological Measurement 1986, Nr 66, str. 629-632.

Hays W. L., *Statistics*, New York 1981.

Hedges L. V. i Olkin I., *Statistical methods for meta-analysis*, New York 1985.

Kozlov V. S. i Erlih J. M., *Obščaja teorija statistiki*, Moskva 1975.

Kožuh B., *Statistične obdelave v pedagoških raziskavah*, Ljubljana 2000.

Kožuh B., *Statistične metode v pedagoškem raziskovanju*, Ljubljana 2003.

Kožuh B., *Statystyka*, Częstochowa 2005.

Krysztofiak M., Urbanek D., *Metody statystyczne*, Warszawa 1977.

Malarska A., Mikulska H., *Statystyka stosowana nie tylko przez psychologów i pedagogów*, Łódź 2000.

Milligan G. W., *A computer program for calculating power of the chi-square test*, Educational and Psychological Measurement 1979, Nr. 39, str. 681-684.

Mužić V., *Uvod u metodologiju istraživanja odgoja i obrazovanja*, Zagreb 1999.

Oakes M., *Statistical inference: A commentary for the social and behavioral sciences*, New York 1986.

Pavkov T. W., Kent A. P., *Do biegu, gotowi – start! Wprowadzenie do SPSS dla Windows*, Gdańsk 2005.

Rao C. R., *Linear statistical inference and its applications*, New York 1975.

Siegel S. i Castellan N. J., *Nonparametric statistics for the behavioral sciences*, New York 1988.

Sobczyk M., *Statystyka*, Warszawa 1995.

Tukey J. W., *Exploratory Data Analysis*, Reading 1977.

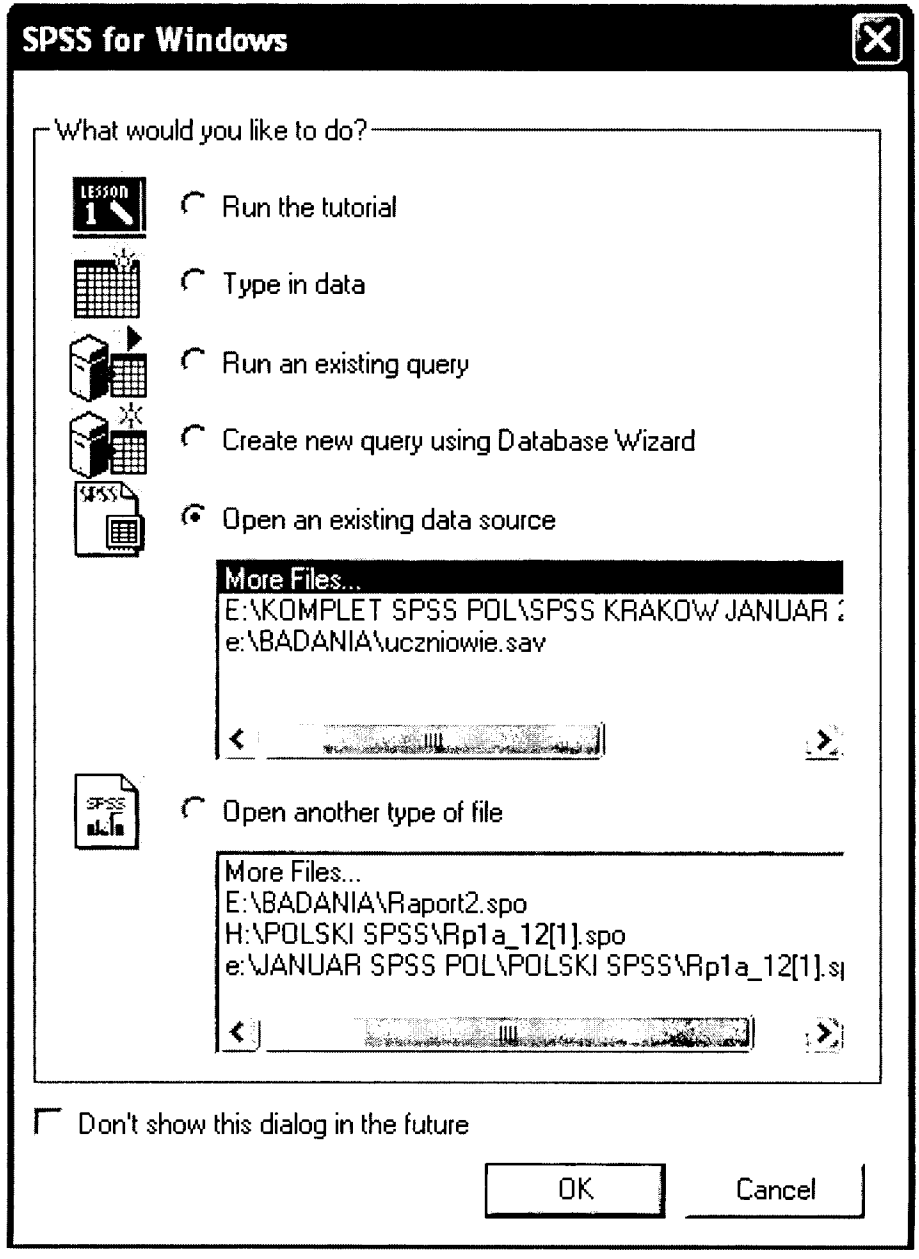
Welkowitz J., Ewen R. B. i Cohen J., *Introductory statistics for the behavioral sciences*, New York 1982.

Winner B.J., *Statistical Principles in Experimental Design*, London 1971.

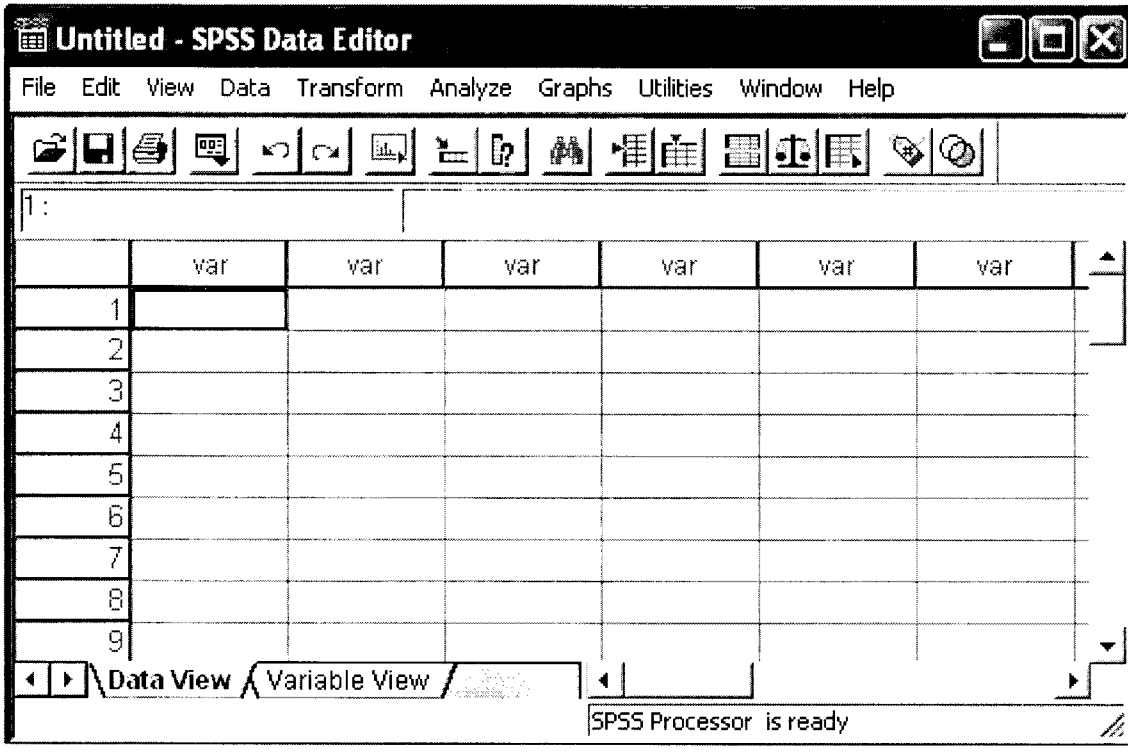
Zaczyński W. P., *Statystyka w pracy badawczej nauczyciela*, Warszawa 1997.

# *Aneksy*

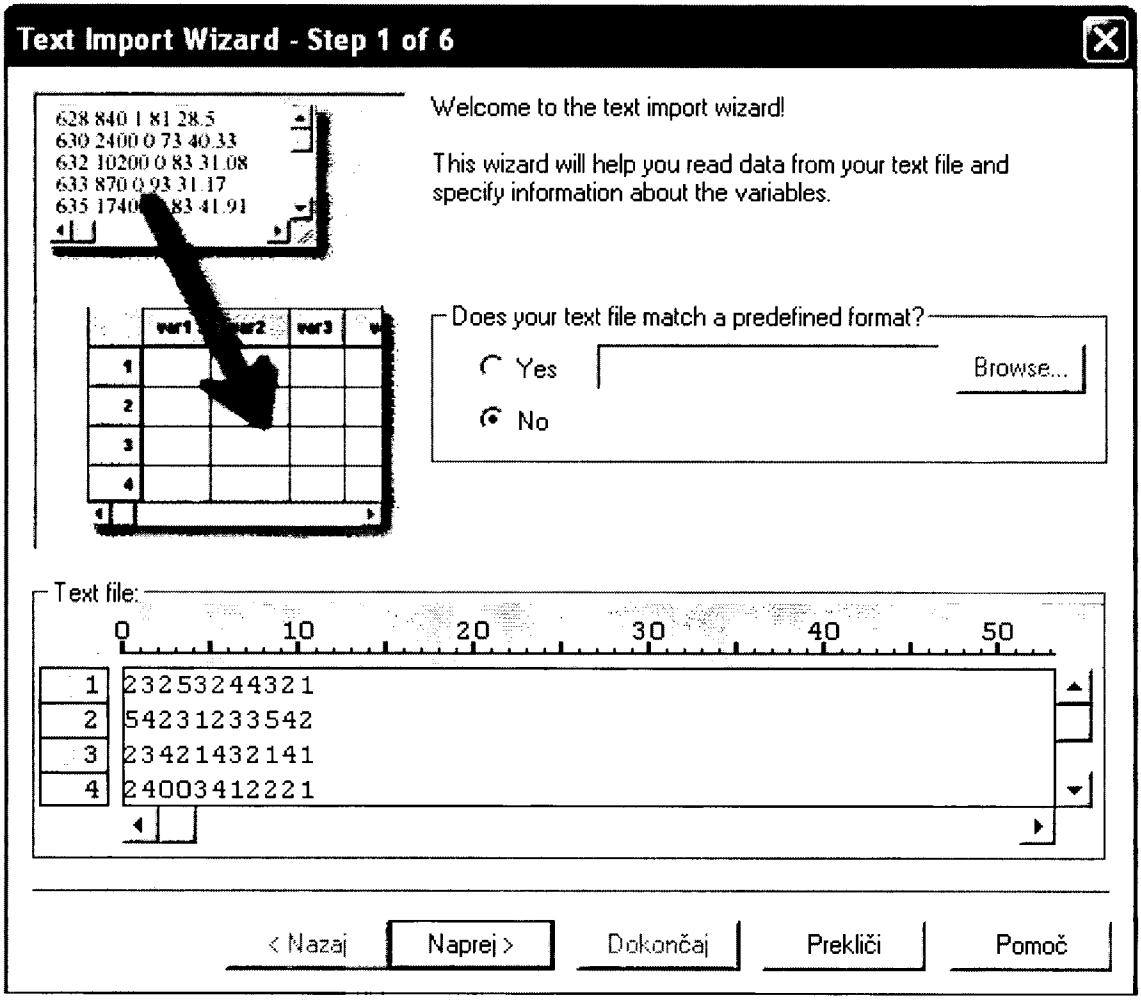
## *Aneks 1. Okno powitalne programu SPSS*



## *Aneks 2. Okno Edytor danych*

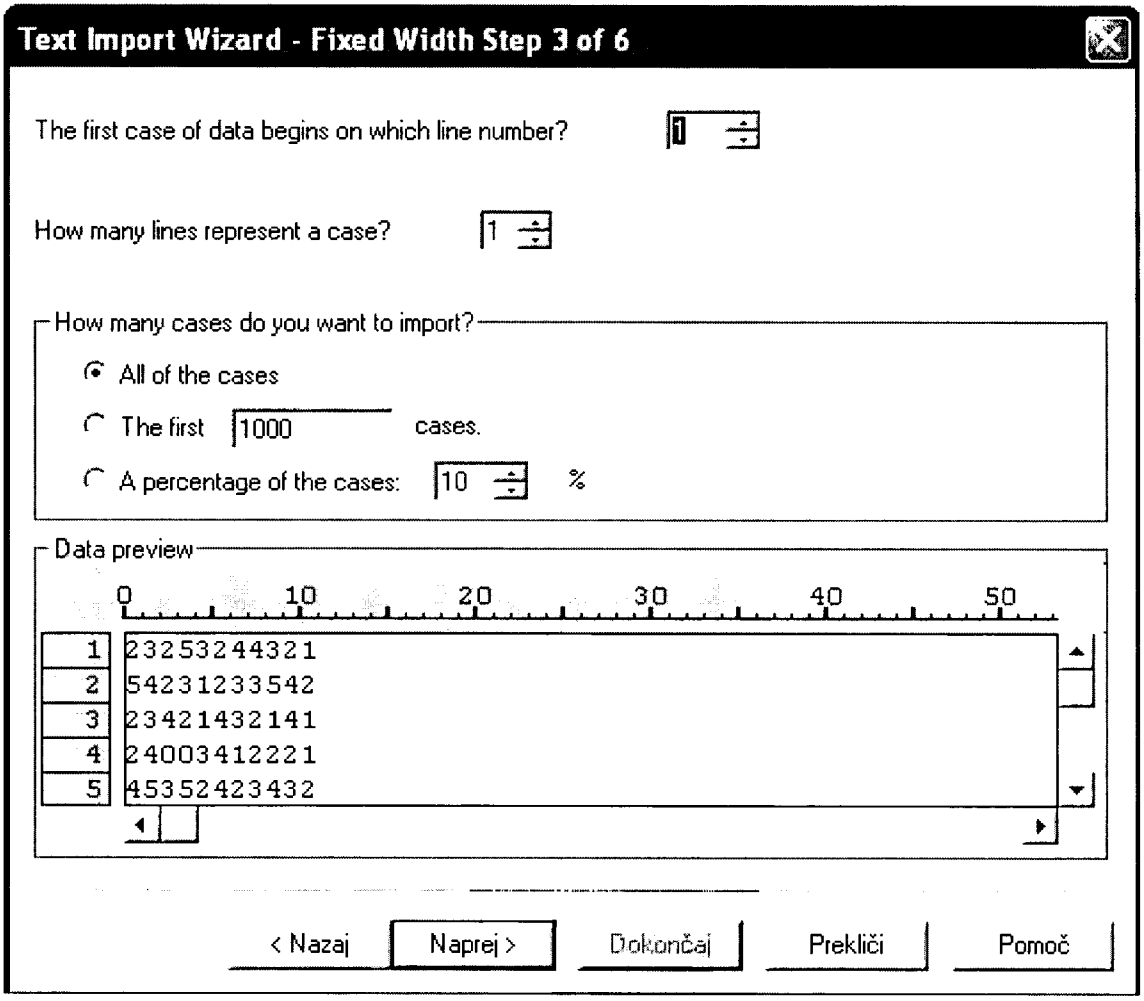


## *Aneks 3. Okno Kreator importu tekstu – krok 1 z 6*

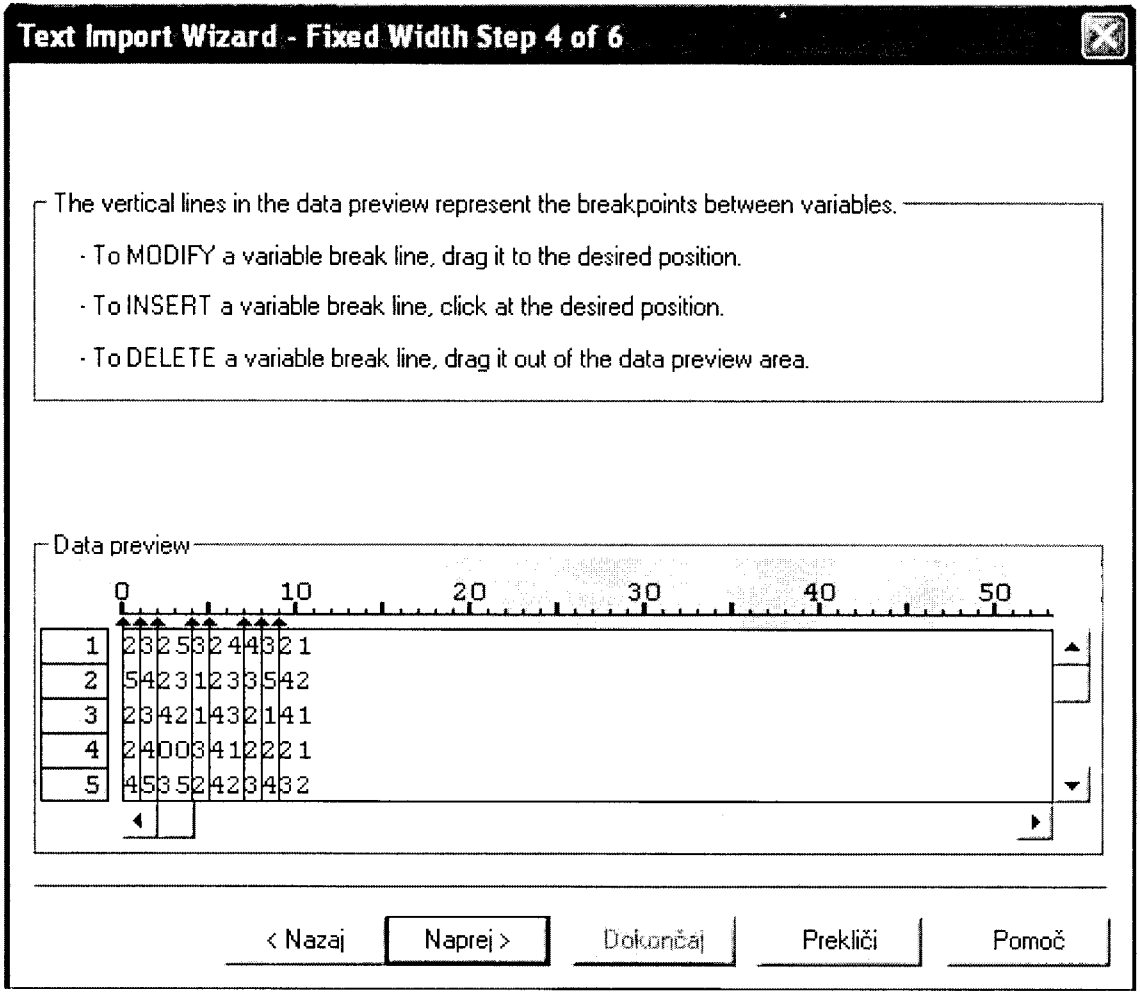


# *Aneksy*

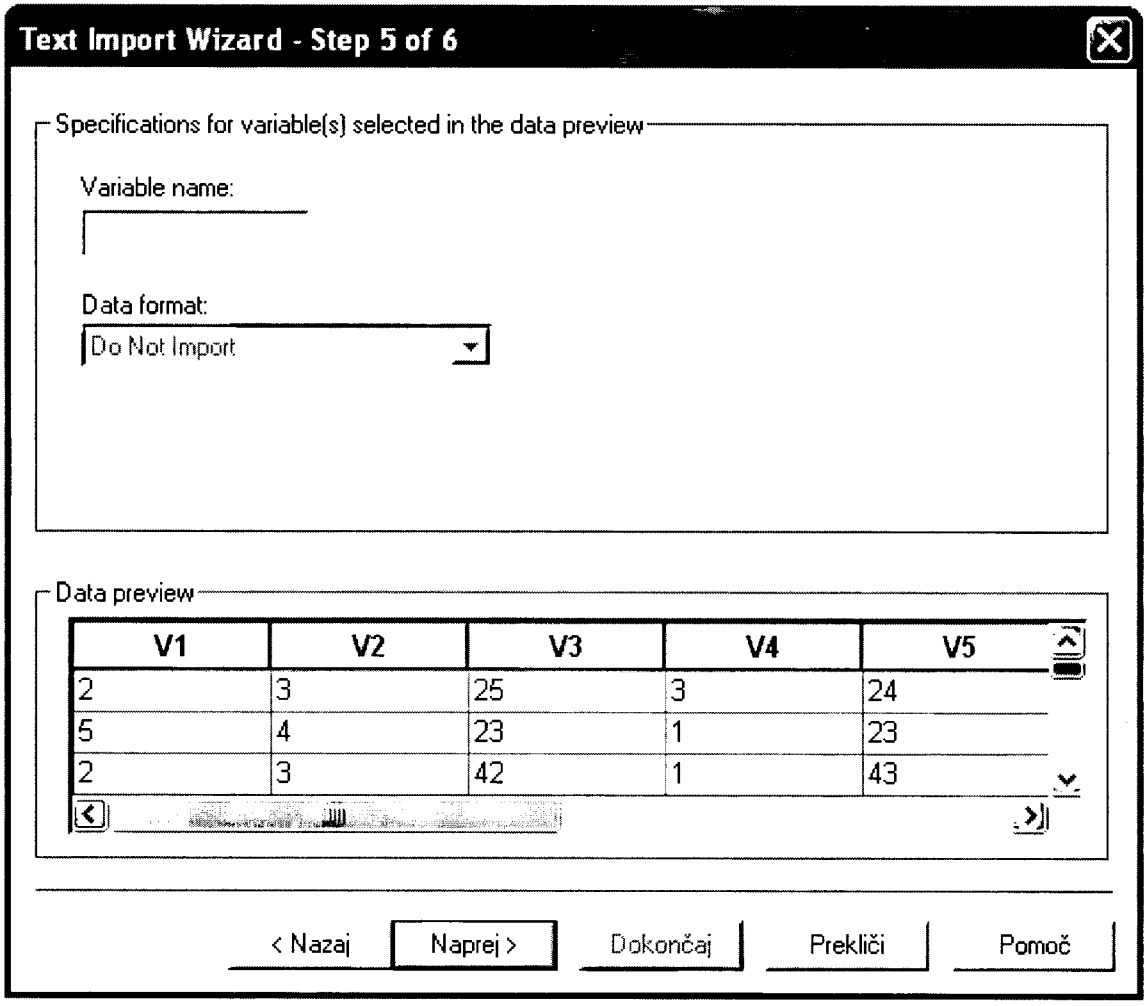
## *Aneks 5. Okno Kreator importu tekstu – krok 3 z 6*



## *Aneks 6. Okno Kreator importu tekstu – krok 4 z 6*

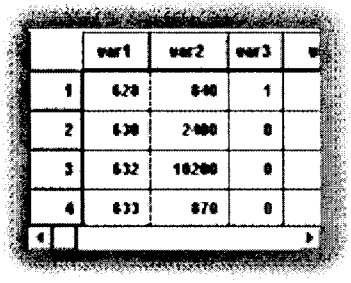


## *Aneks 7. Okno Kreator importu tekstu – krok 5 z 6*



## *Aneks 8. Okno Kreator importu tekstu – krok 6 z 6*

**Text Import Wizard - Step 6 of 6**



You have successfully defined the format of your text file.

Would you like to save this file format for future use?

- Yes — Save As...
- No

Would you like to paste the syntax?

- Yes
- No

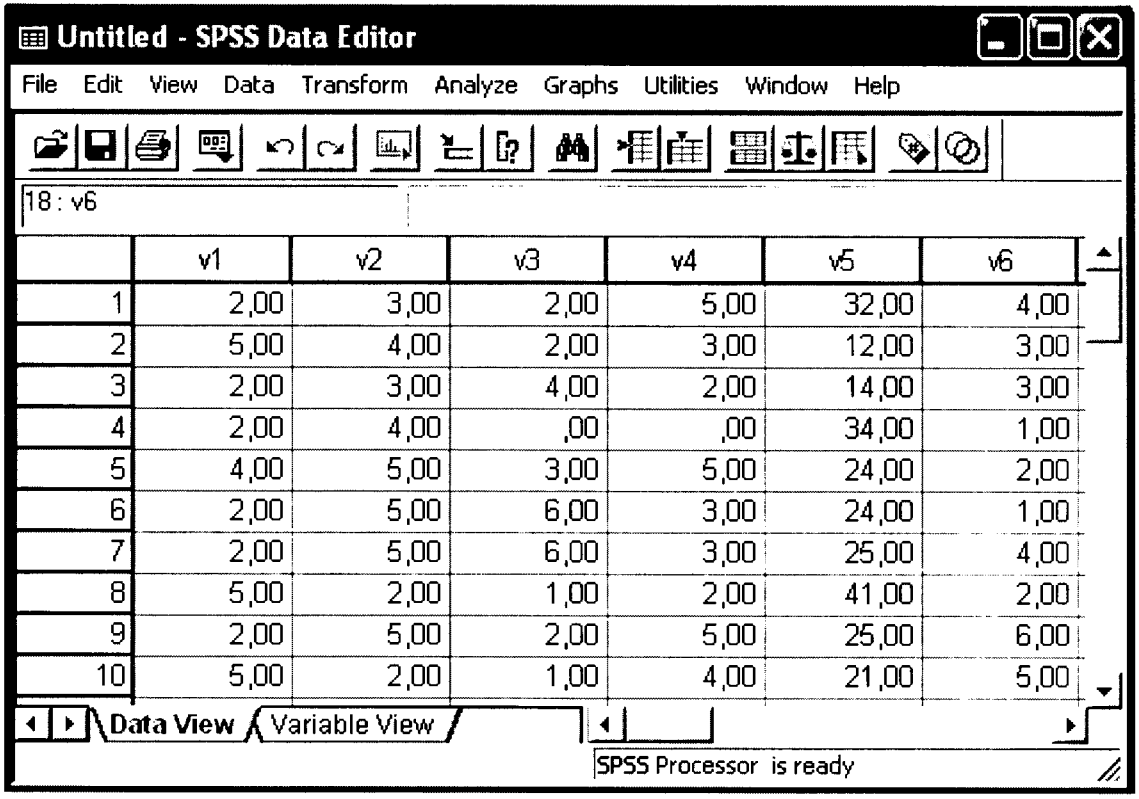
Press the Finish button to complete the text import wizard.

Data preview

| V1 | V2 | V3 | V4 | V5 |
|---|---|---|---|---|
| 2 | 3 | 25 | 3 | 24 |
| 5 | 4 | 23 | 1 | 23 |
| 2 | 3 | 42 | 1 | 43 |

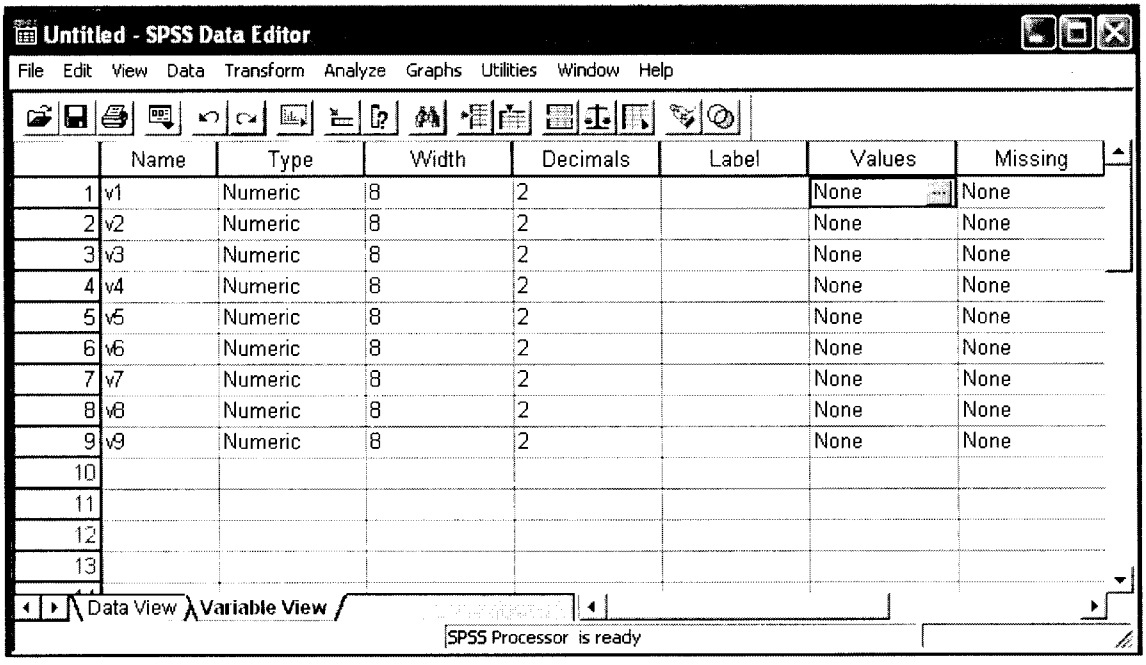
< Nazaj | Naprej > | Dokončaj | Prekliči | Pomoč

## *Aneks 9. Okno Edytor danych – widok na dane*

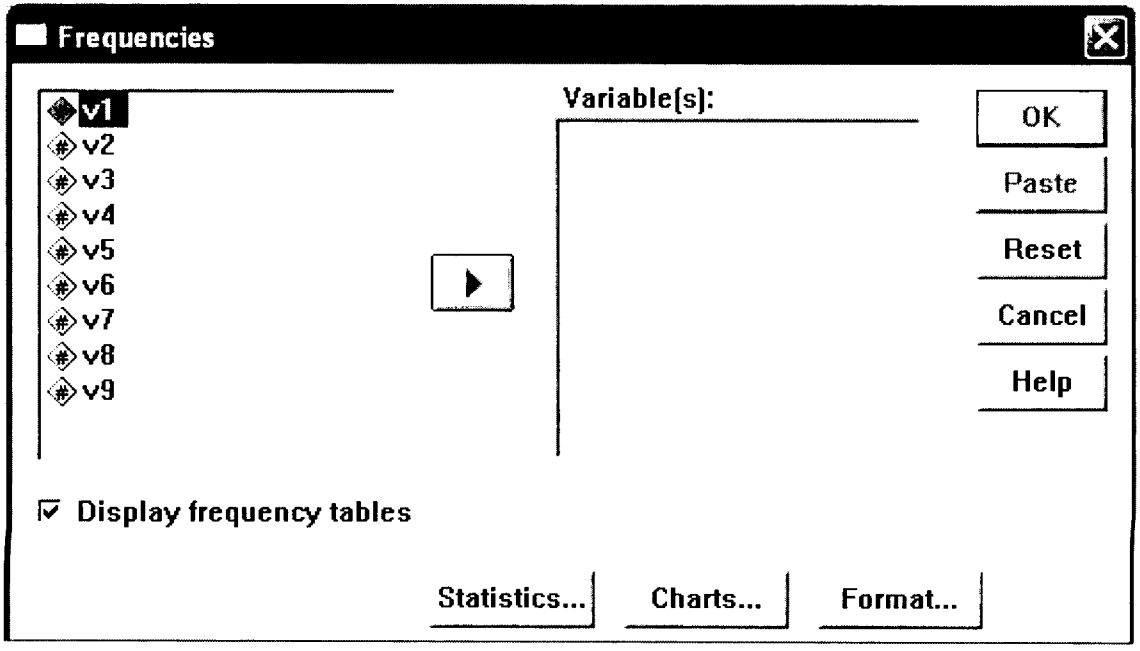



| | v1 | v2 | v3 | v4 | v5 | v6 |
|---|---|---|---|---|---|---|
| 1 | 2,00 | 3,00 | 2,00 | 5,00 | 32,00 | 4,00 |
| 2 | 5,00 | 4,00 | 2,00 | 3,00 | 12,00 | 3,00 |
| 3 | 2,00 | 3,00 | 4,00 | 2,00 | 14,00 | 3,00 |
| 4 | 2,00 | 4,00 | ,00 | ,00 | 34,00 | 1,00 |
| 5 | 4,00 | 5,00 | 3,00 | 5,00 | 24,00 | 2,00 |
| 6 | 2,00 | 5,00 | 6,00 | 3,00 | 24,00 | 1,00 |
| 7 | 2,00 | 5,00 | 6,00 | 3,00 | 25,00 | 4,00 |
| 8 | 5,00 | 2,00 | 1,00 | 2,00 | 41,00 | 2,00 |
| 9 | 2,00 | 5,00 | 2,00 | 5,00 | 25,00 | 6,00 |
| 10 | 5,00 | 2,00 | 1,00 | 4,00 | 21,00 | 5,00 |

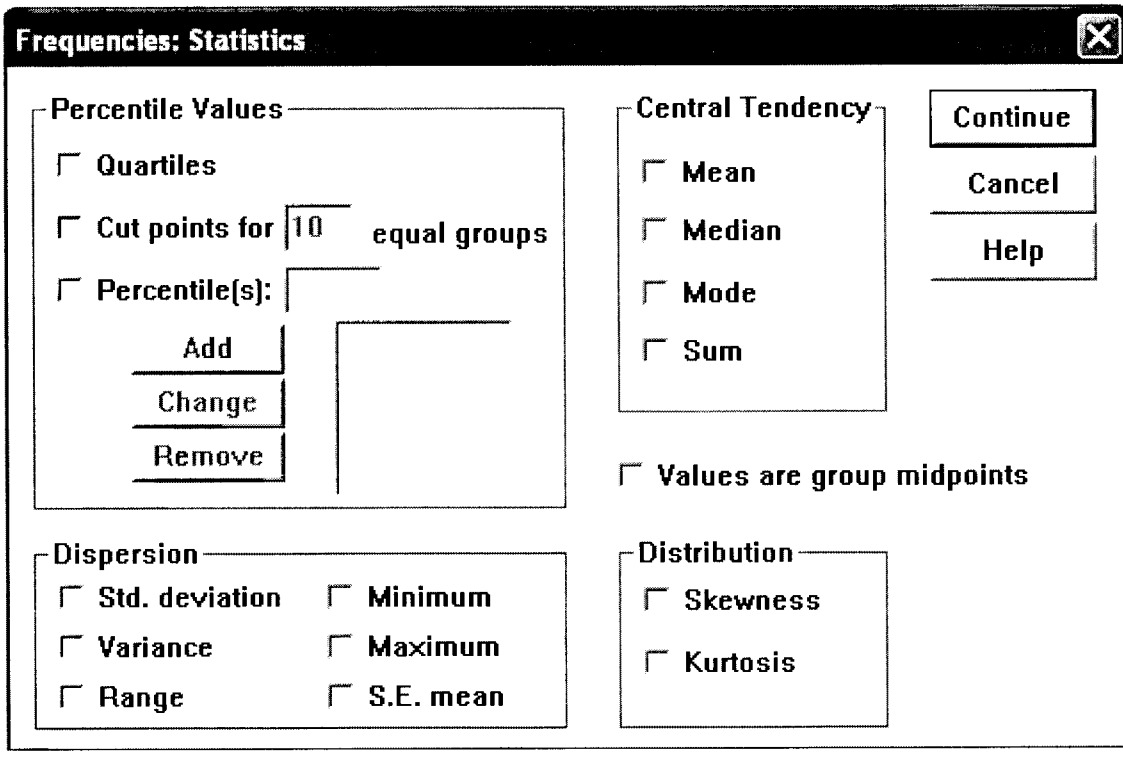
# *Aneks 10. Okno Edytor danych – widok na zmienne*



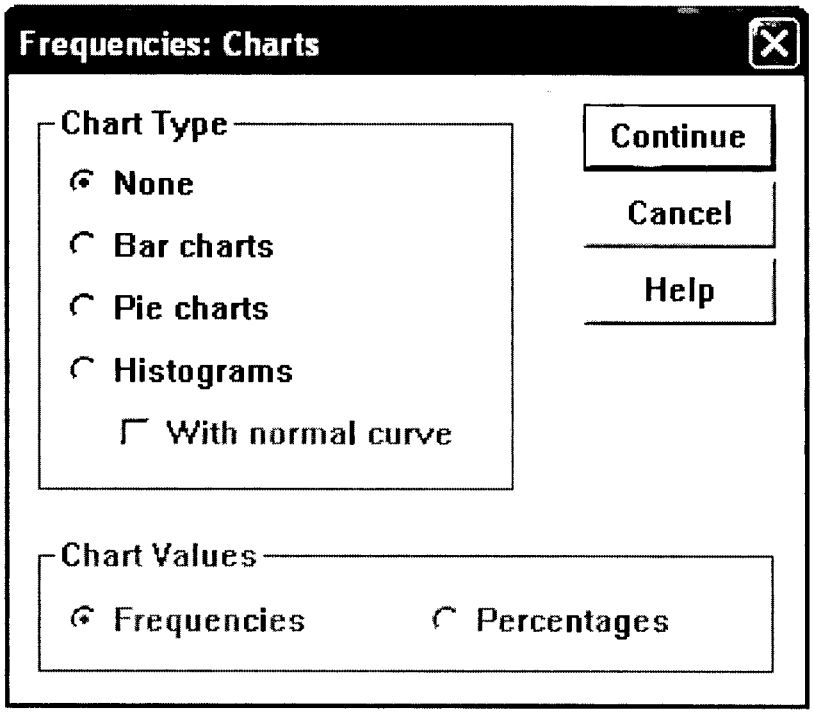
## *Aneks 11. Okno Częstości*



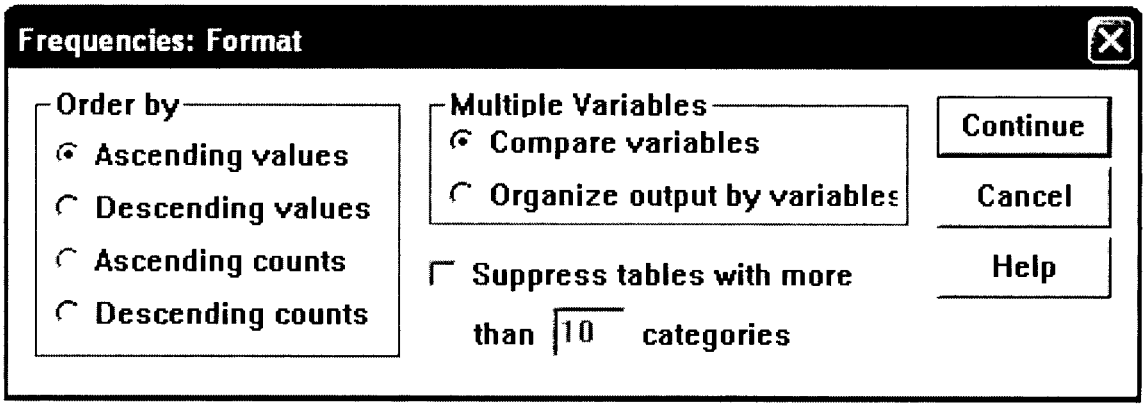
## *Aneks 12. Okno Częstości: Statystyki*



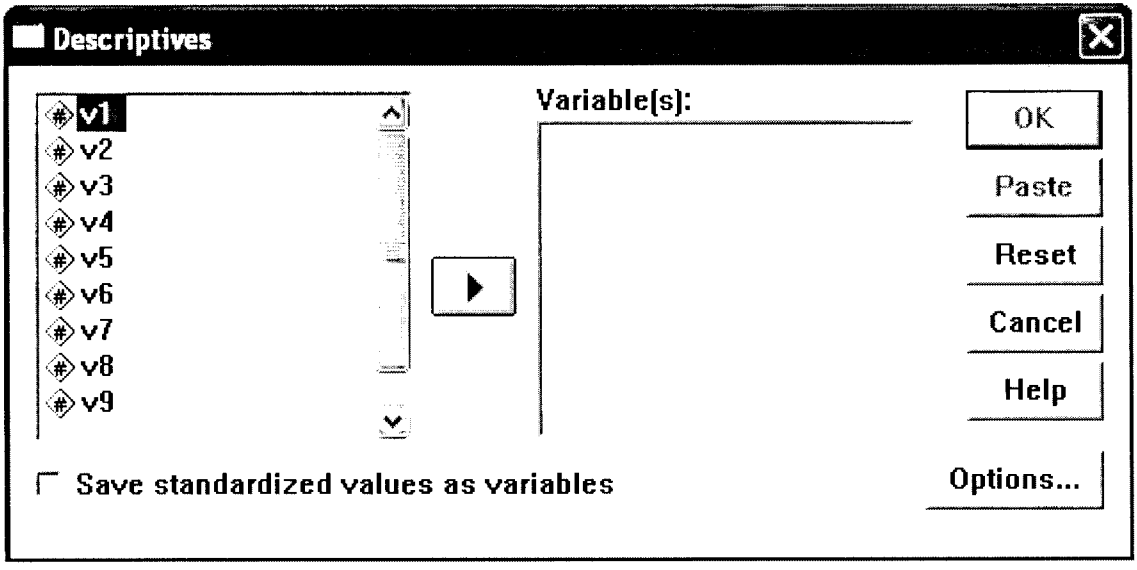
*Aneks 13. Okno Częstości: Wykresy*



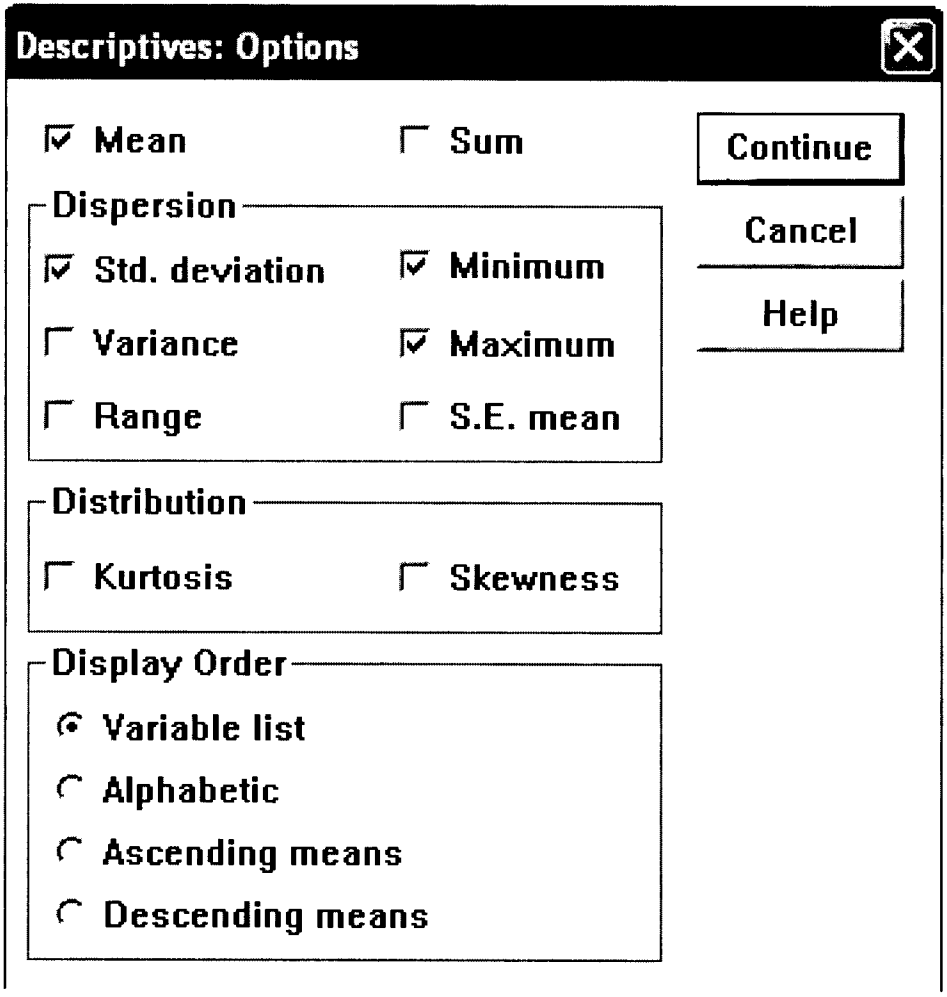

*Aneks 14. Okno Częstości: Format*



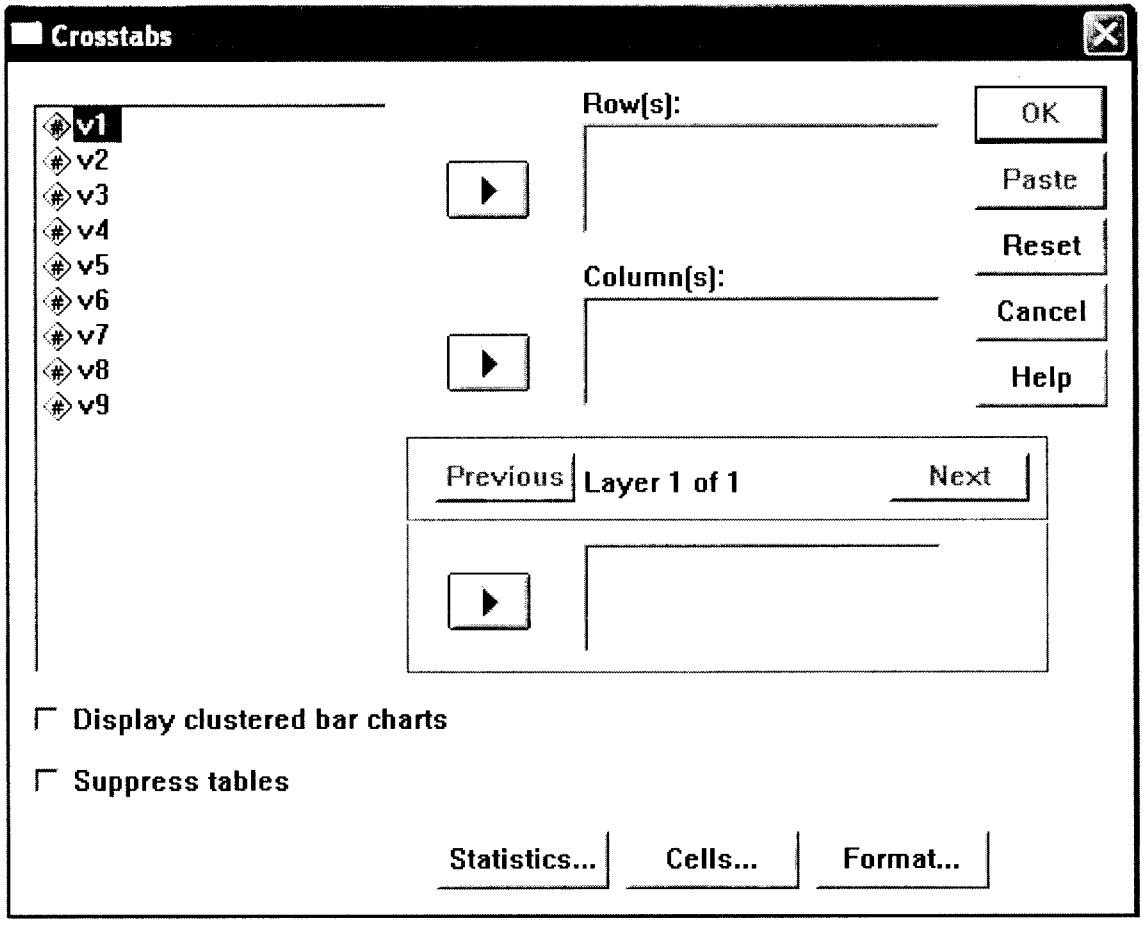
## *Aneks 15. Okno Statystyki opisowe*



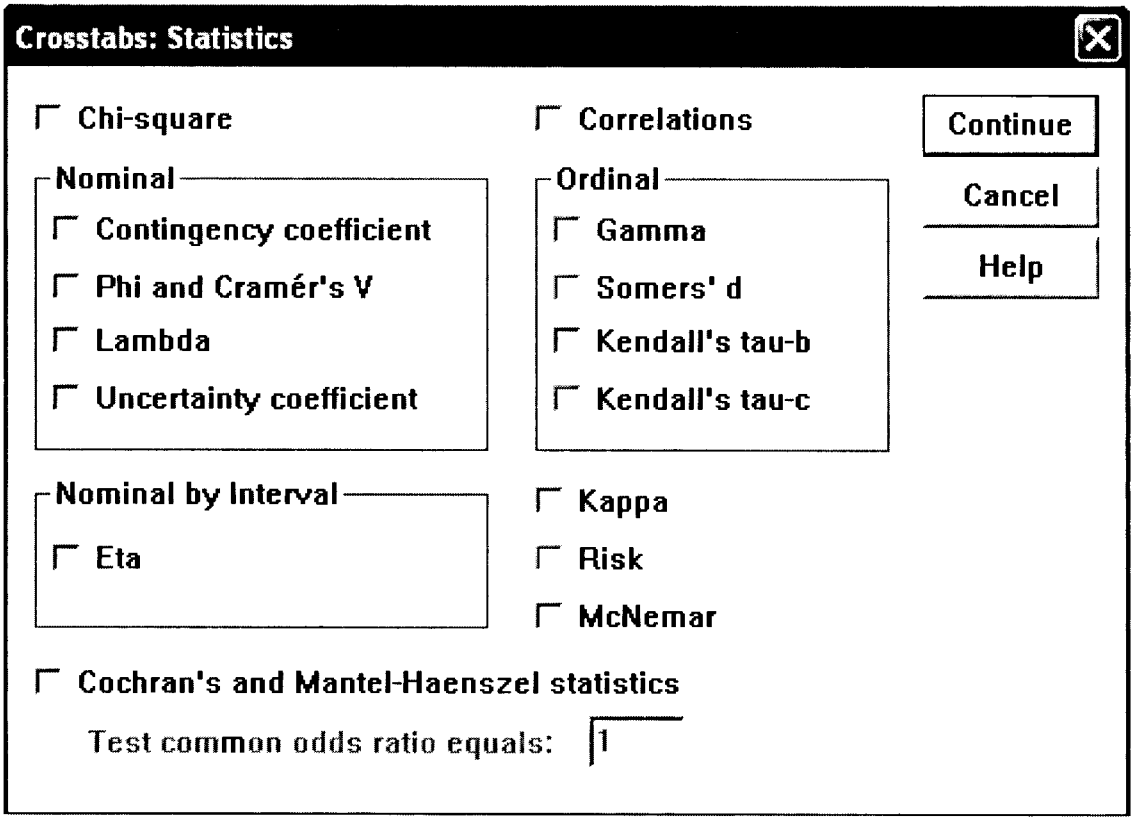
*Aneks 16. Okno Statystyki opisowe: Opcje*



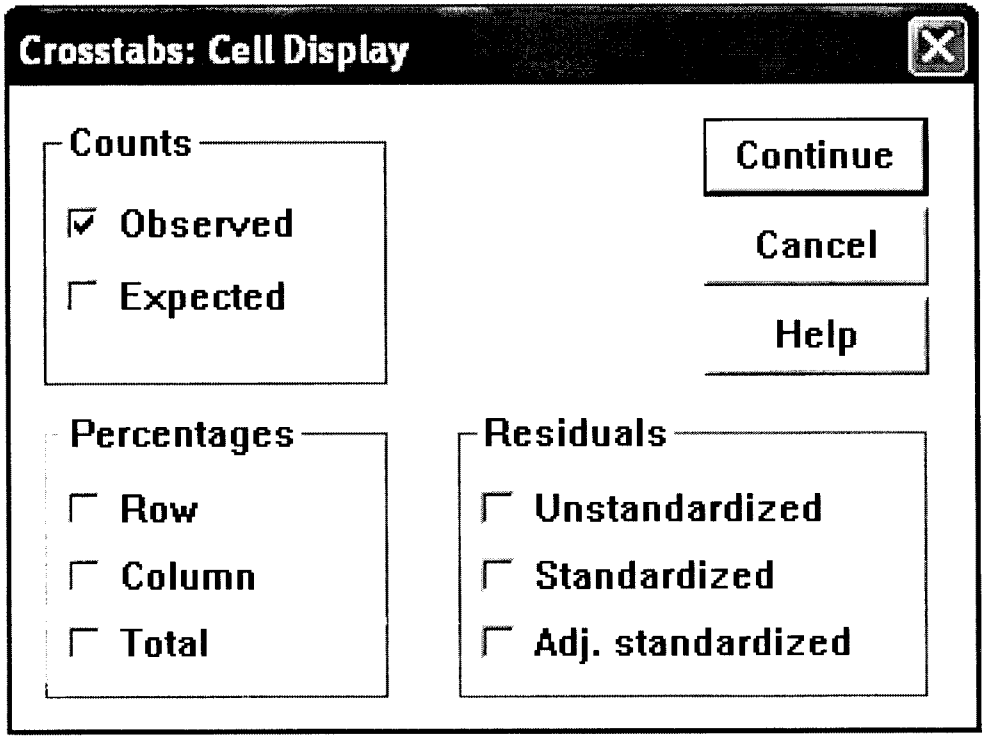
## *Aneks 17. Okno Tabele krzyżowe*



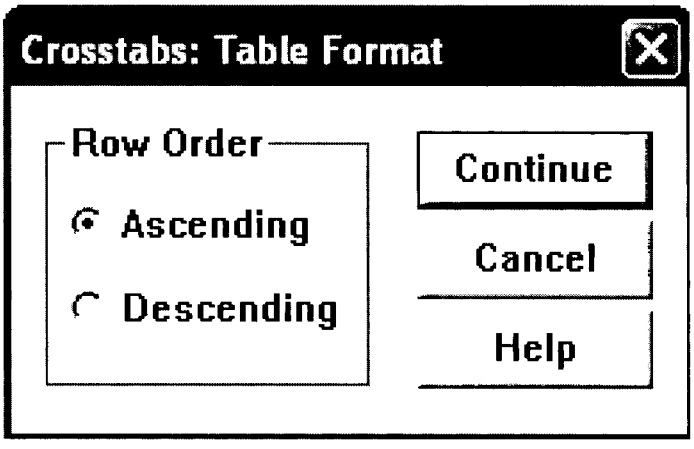
*Aneks 18. Okno Tabele krzyżowe: Statystyki*



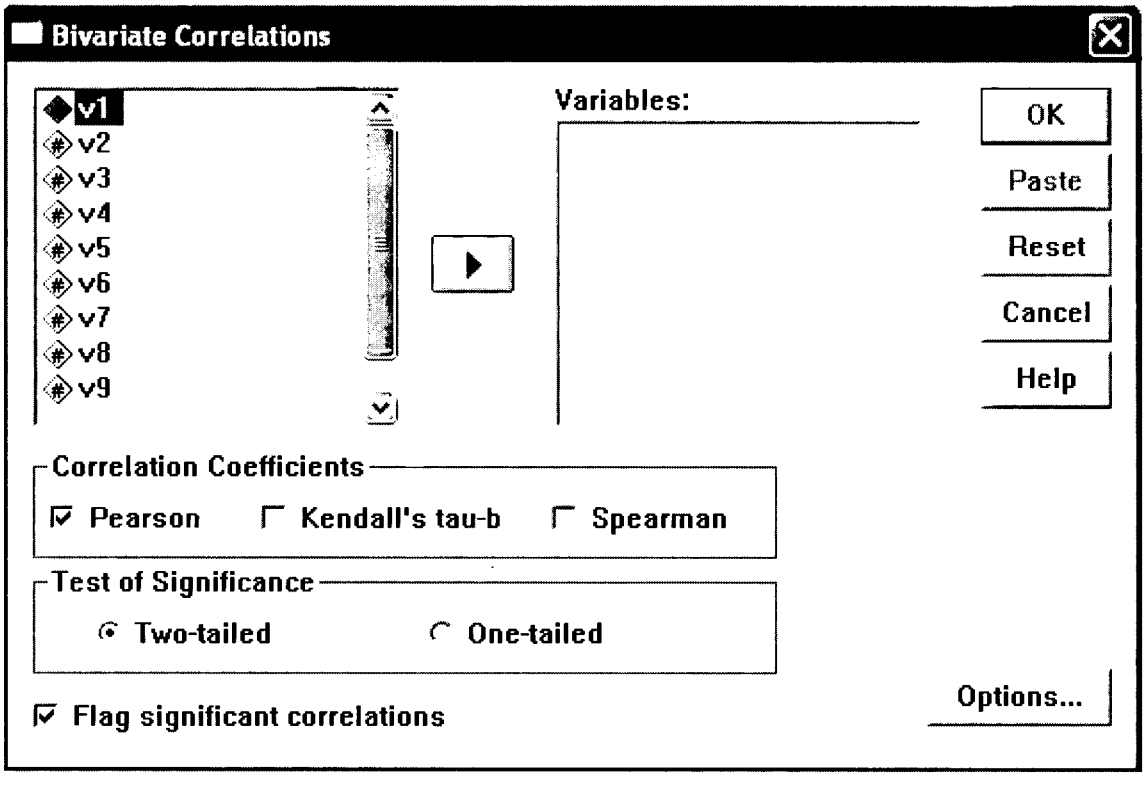
*Aneks 19. Okno Tabele krzyżowe: Zawartość komórek*



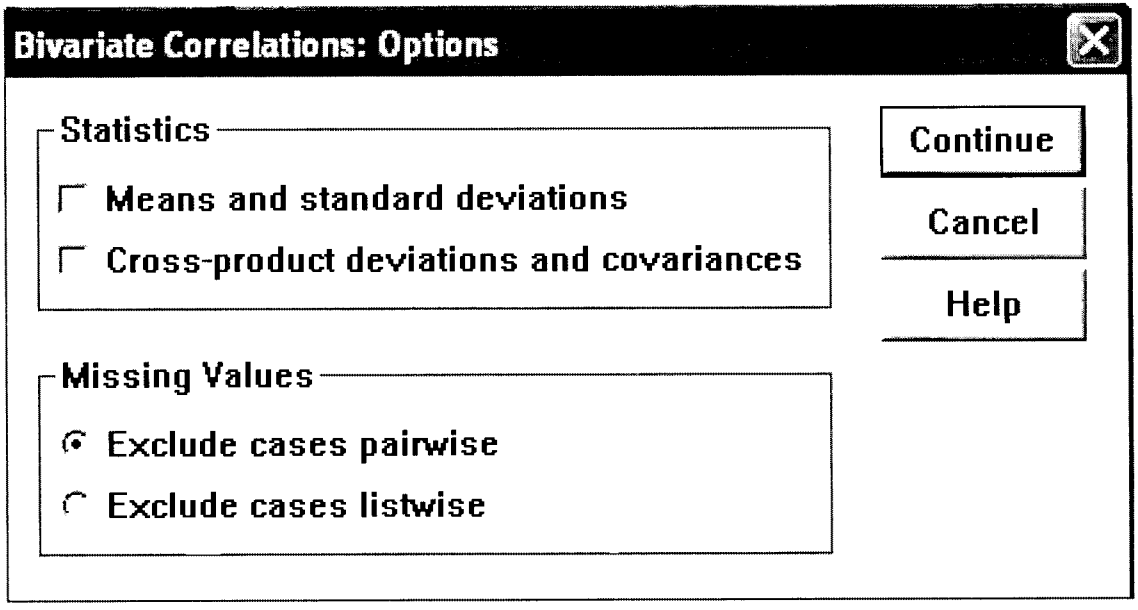
*Aneks 20. Okno Tabele krzyżowe: Format*



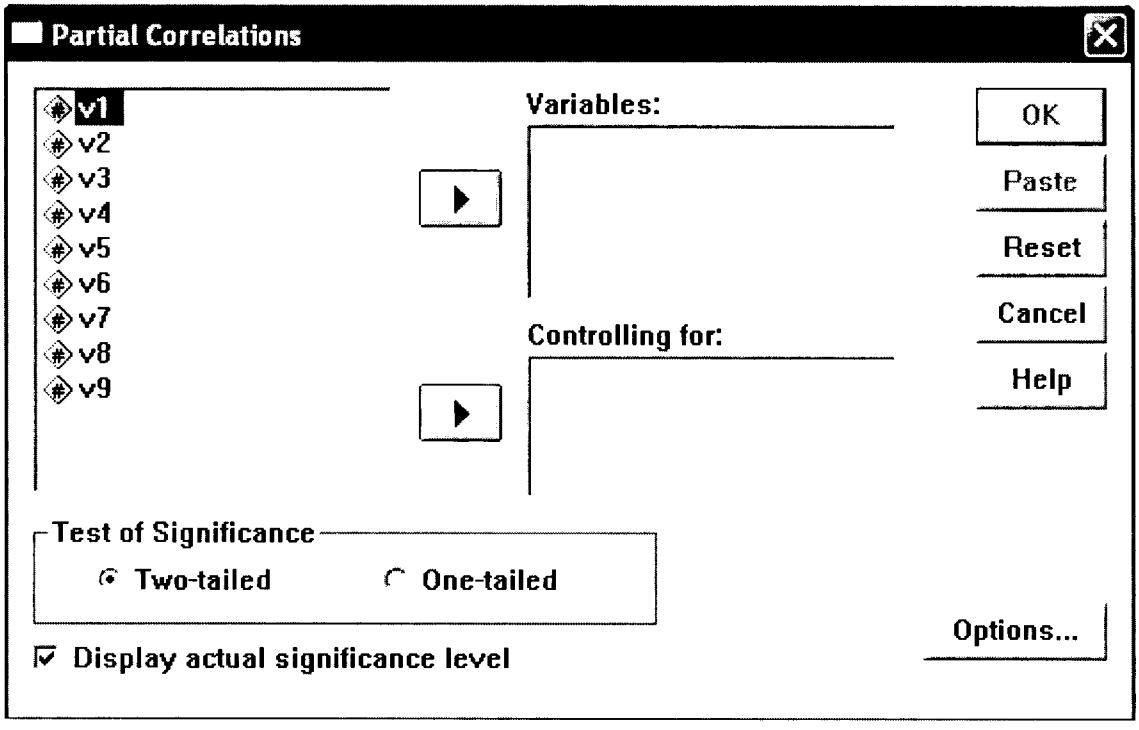

*Aneks 21. Okno Korelacje parami*



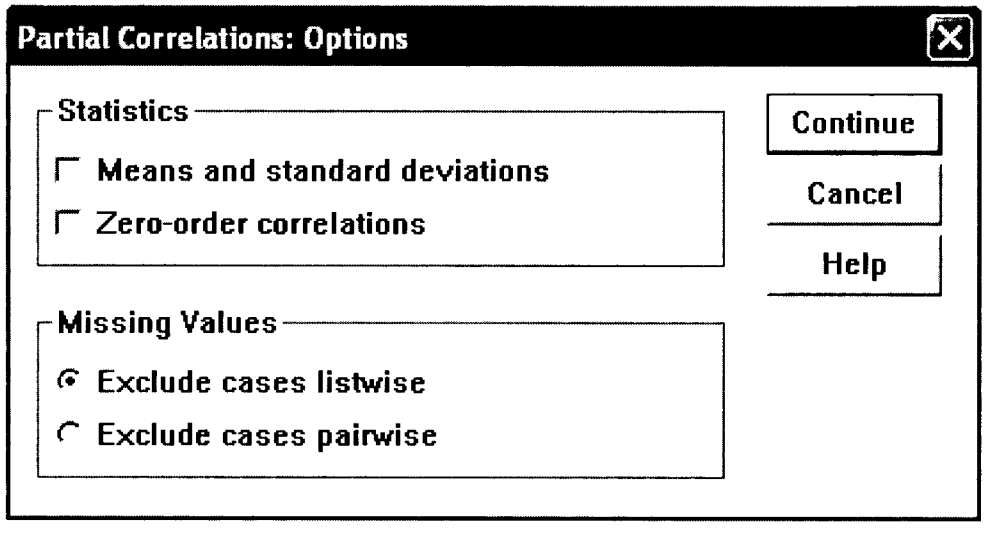
*Aneks 22. Okno Korelacje parami: Opcje*



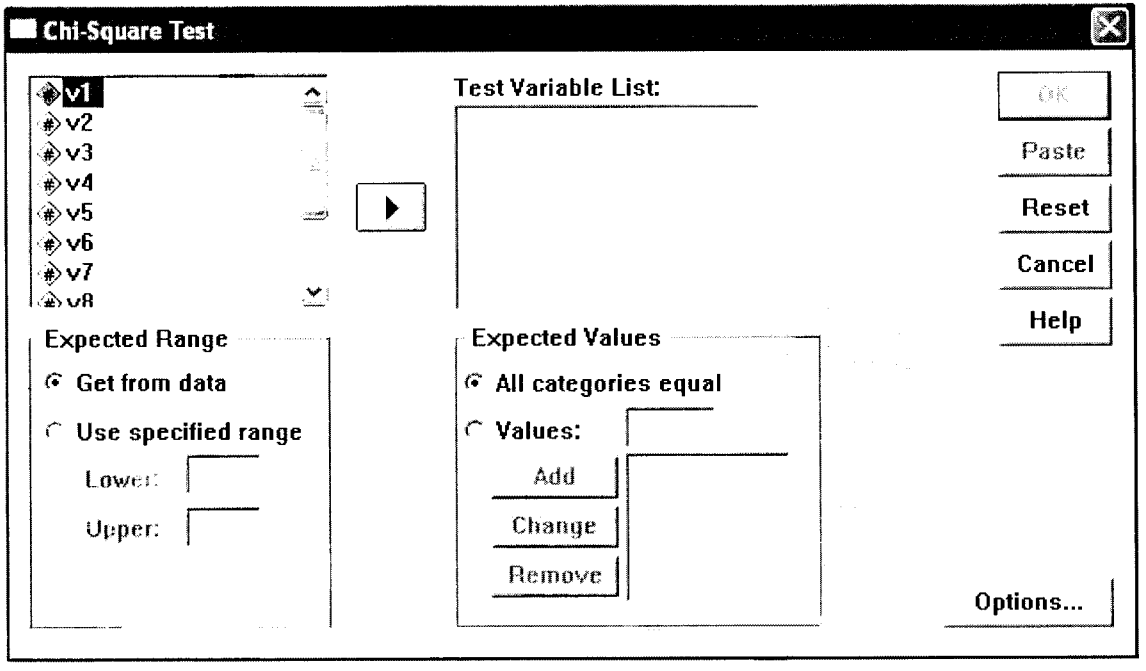
*Aneks 23. Okno Korelacje cząstkowe*



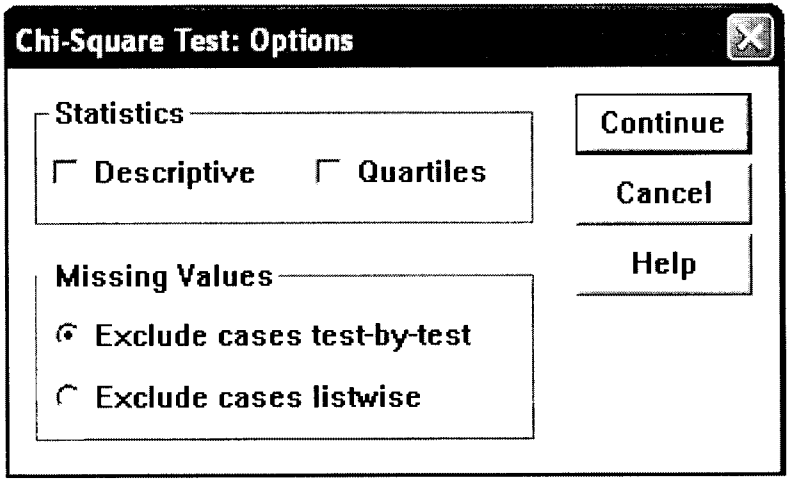

*Aneks 24. Okno Korelacje cząstkowe: Opcje*



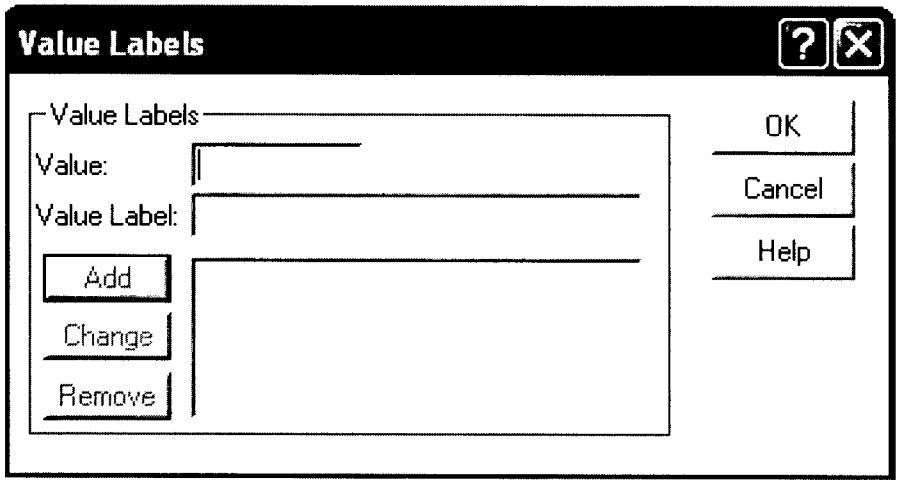
*Aneks 25. Okno Test chi-kwadrat*



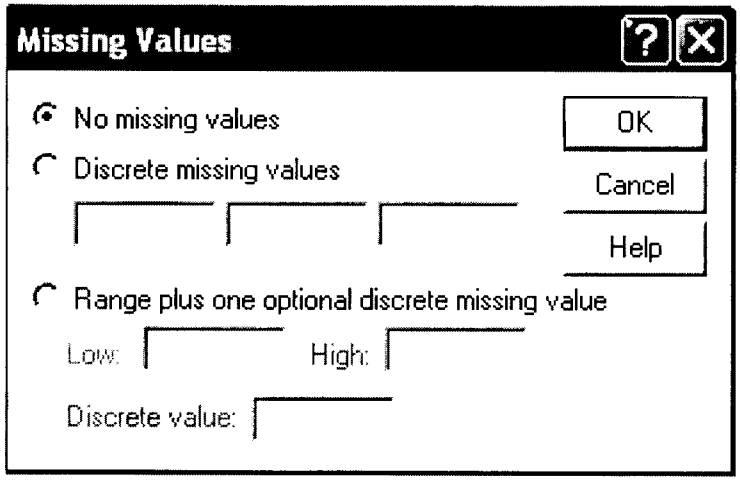

*Aneks 26. Okno Test chi-kwadrat: Opcje*



## *Aneks 27. Okno Wartości*




## *Aneks 28. Okno Braki danych*



## *Aneks 29. Okno Edytor poleceń*