

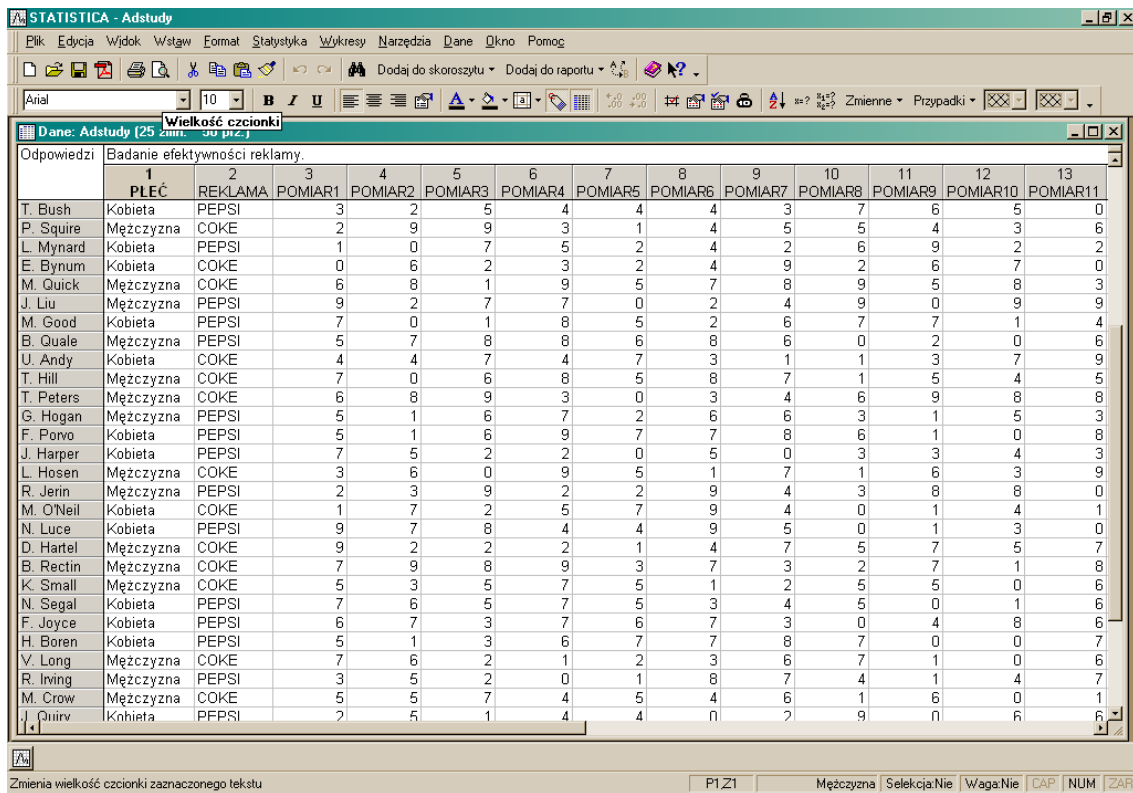
Ćwiczenie 2: Prosty projekt tworzący standardowe raporty i podsumowania

Plik danych *Adstudy.sta* zawiera:

- nazwiska (w kolumnie z nazwami przypadków) osób oceniających skuteczność reklamy napoi orzeźwiających
- w kolumnie *pleć* odnotowana jest płeć osoby
- kolumna *Reklama* zawiera nazwę ocenianego napoju *Coke* lub *Pepsi*
- w zmiennych od *Pomiar1* do *Pomiar23* zawarta jest ocena (w skali od 1 do 9; 0 oznacza brak zdania) skuteczności reklamy (z ostatnich 23 miesięcy), przeprowadzonej przez pewną firmę.

Cel analiz: utworzenie projektu data mining, który umożliwi szybkie wskazanie:

- osób najwyżej oceniających reklamę oraz
- skuteczności reklamy z okresu 3 miesięcy i 6 miesięcy (i porównanie czy zmieniała się w czasie), z uwzględnieniem podziału na dwa badane produkty *Coke* i *Pepsi*



The screenshot shows the STATISTICA - Adstudy interface. The main window displays a data table titled 'Dane: Adstudy (25 zm., 30 prz.)'. The table has 14 columns: '1 PLEĆ', '2 REKLAMA', '3 POMIAR1', '4 POMIAR2', '5 POMIAR3', '6 POMIAR4', '7 POMIAR5', '8 POMIAR6', '9 POMIAR7', '10 POMIAR8', '11 POMIAR9', '12 POMIAR10', '13 POMIAR11', and '14'. The rows list individuals and their corresponding data values. The bottom status bar shows 'Zmienia wielkość czcionki zaznaczonego tekstu' and 'P1.Z1 Męczyzna Selekcja:Nie Waga:Nie CAP NUM ZAR'.

	1 PLEĆ	2 REKLAMA	3 POMIAR1	4 POMIAR2	5 POMIAR3	6 POMIAR4	7 POMIAR5	8 POMIAR6	9 POMIAR7	10 POMIAR8	11 POMIAR9	12 POMIAR10	13 POMIAR11	14
T. Bush	Kobieta	PEPSI	3	2	5	4	4	4	3	7	6	5	0	
P. Squire	Męczyzna	COKE	2	9	9	3	1	4	5	5	4	3	6	
L. Mynard	Kobieta	PEPSI	1	0	7	5	2	4	2	6	9	2	2	
E. Bynum	Kobieta	COKE	0	6	2	3	2	4	9	2	6	7	0	
M. Quick	Męczyzna	COKE	6	8	1	9	5	7	8	9	5	8	3	
J. Liu	Męczyzna	PEPSI	9	2	7	7	0	2	4	9	0	9	9	
M. Good	Kobieta	PEPSI	7	0	1	8	5	2	6	7	7	1	4	
B. Quale	Męczyzna	PEPSI	5	7	8	8	6	8	6	0	2	0	6	
U. Andy	Kobieta	COKE	4	4	7	4	7	3	1	1	3	7	9	
T. Hill	Męczyzna	COKE	7	0	6	8	5	8	7	1	5	4	5	
T. Peters	Męczyzna	COKE	6	8	9	3	0	3	4	6	9	8	8	
G. Hogan	Męczyzna	PEPSI	5	1	6	7	2	6	6	3	1	5	3	
F. Porvo	Kobieta	PEPSI	5	1	6	9	7	7	8	6	1	0	8	
J. Harper	Kobieta	PEPSI	7	5	2	2	0	5	0	3	3	4	3	
L. Hosen	Męczyzna	COKE	3	6	0	9	5	1	7	1	6	3	9	
R. Jerin	Męczyzna	PEPSI	2	3	9	2	2	9	4	3	8	8	0	
M. O'Neil	Kobieta	COKE	1	7	2	5	7	9	4	0	1	4	1	
N. Luce	Kobieta	PEPSI	9	7	8	4	4	9	5	0	1	3	0	
D. Hartel	Męczyzna	COKE	9	2	2	2	1	4	7	5	7	5	7	
B. Rectin	Męczyzna	COKE	7	9	8	9	3	7	3	2	7	1	8	
K. Small	Męczyzna	COKE	5	3	5	7	5	1	2	5	5	0	6	
N. Segal	Kobieta	PEPSI	7	6	5	7	5	3	4	5	0	1	6	
F. Joyce	Kobieta	PEPSI	6	7	3	7	6	7	3	0	4	8	6	
H. Boren	Kobieta	PEPSI	5	1	3	6	7	7	8	7	0	0	7	
V. Long	Męczyzna	COKE	7	6	2	1	2	3	6	7	1	0	6	
R. Irving	Męczyzna	PEPSI	3	5	2	0	1	8	7	4	1	4	7	
M. Crow	Męczyzna	COKE	5	5	7	4	5	4	6	1	6	0	1	
J. Onirv	Kobieta	PEPSI	2	5	1	4	4	0	2	9	0	6	6	

1. Otwórz dokument
2. Otwórz przestrzeń roboczą i wstaw źródło
3. Otwórz przeglądarkę węzłów
4. Wstaw węzeł Dane->Przekształcenia->Czyszczenie -> Czyszczenie danych
5. Ustaw parametry węzła:
 - a. wybierz zmienne na wejściu: POMIAR1-POMIAR23
 - b. na wyjściu: wszystkie

- c. „Ustawienia przetwarzania braków danych” wybierz wszystkie zmienne, w „Działanie” wybierz „Zastąp BD podaną wartością” i w „Dodatkowe kody BD” wpisz wszędzie „0”

Zamiana braków danych

Wejście:

Zmienne: POMIAR1-POMIAR23

Przypadki: WSZYSTKIE

☐ Stosuj wagi

Wartości dodatkowo traktowane jako braki danych:

☐ Inne niż liczby ☐ Niewidoczne znaki

Ustawienia przetwarzania braków danych

Zmienna	Działanie	Dodatkowe kody BD	Zamień na	Próg wyłączenia
POMIAR1	Zastąp BD podaną wartością	0		0
POMIAR2	Zastąp BD podaną wartością	0		0
POMIAR3	Zastąp BD podaną wartością	0		0
POMIAR4	Zastąp BD podaną wartością	0		0
POMIAR5	Zastąp BD podaną wartością	0		0
POMIAR6	Zastąp BD podaną wartością	0		0
POMIAR7	Zastąp BD podaną wartością	0		0
POMIAR8	Zastąp BD podaną wartością	0		0
POMIAR9	Zastąp BD podaną wartością	0		0

Wyjście

Zmienne: WSZYSTKIE

☐ Usuń zmienne wg kryterium wyłączenia ☐ Skopiuj formaty

Uwaga: Aby wybrać wiele wierszy klikamy je przy wciśniętym klawisz Shift lub Ctrl. Zmiana ustawień dotyczy wszystkich wybranych wierszy. Naciśnięcie OK powoduje zmianę braków danych zgodnie z ustawieniami.

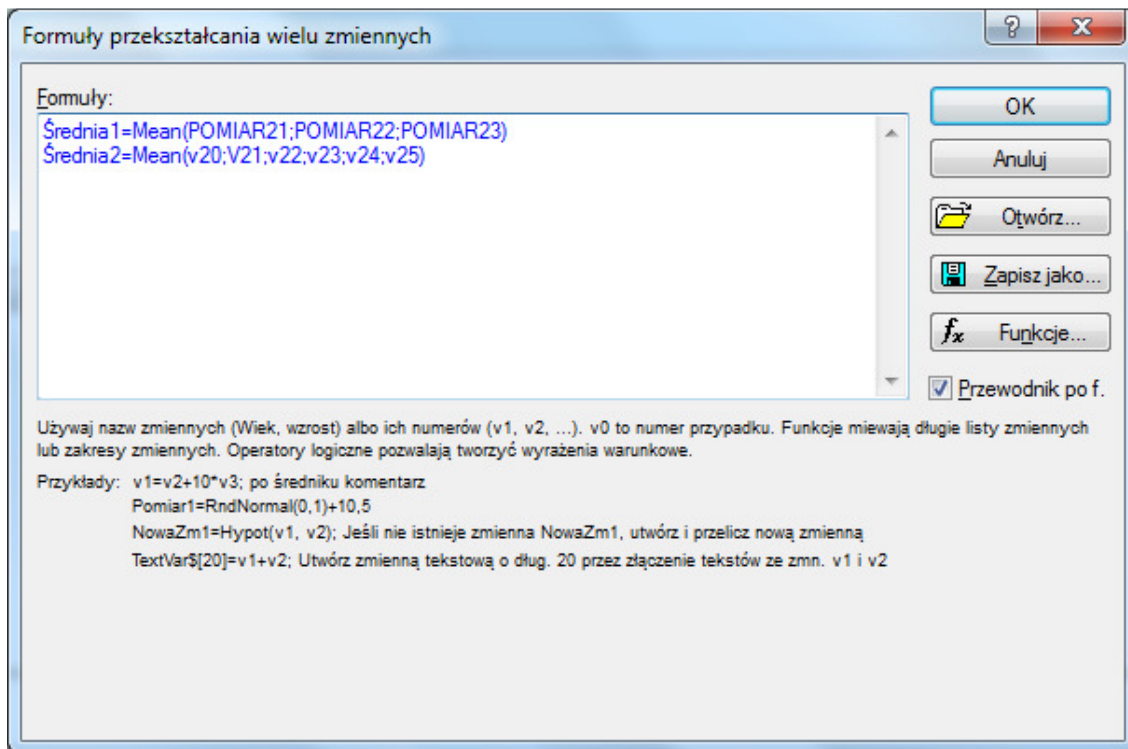
OK Anuluj

6. Po sprawdzeniu powstałego dokumentu widzimy, że wszystkie zera z pliku z danymi zostały usunięte i zastąpione brakami danych.

Transformacja danych.

Obliczmy średnią ocenę w ostatnich trzech oraz sześciu miesiącach, tzn. średnią z wartości zmiennych od *POMIAR21* (zmienna 23) do *POMIAR23* (zmienna 25) oraz od *POMIAR18* (zmienna 20) do *POMIAR23* (zmienna 25).

1. Wstaw węzeł Dane -> Przekształcenia -> Przekształcenia zmiennych
2. Wpisz przekształcenia:

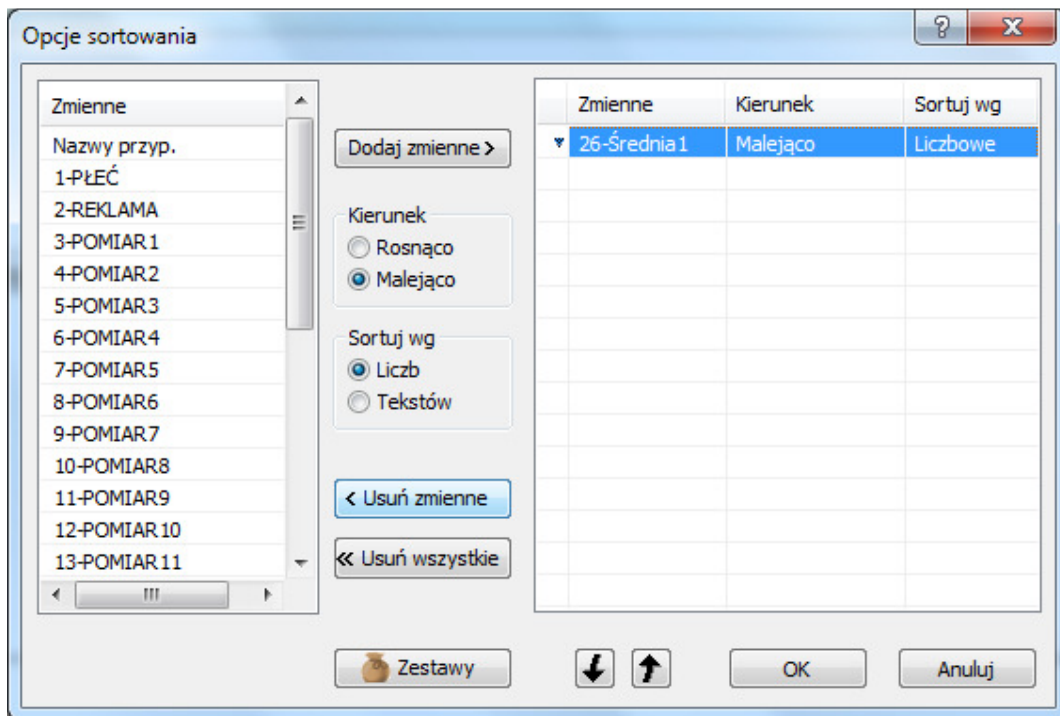


3. Uruchom węzeł i sprawdź, że dodały się nowe zmienne

Sortowanie danych.

Teraz posortujemy dane względem dwóch utworzonych zmiennych, czyli średnich ocen z ostatnich 3 oraz 6 miesięcy.

1. Wstaw 2 węzły sortowania: Dane -> Przypadki -> Sortuj
2. Pierwszemu węzłowi zmień nazwę na „3 mies” a drugiemu na „6 mies”
3. W pierwszym węźle wybierz parametry (Średnia1 to średnia z 3 mies):



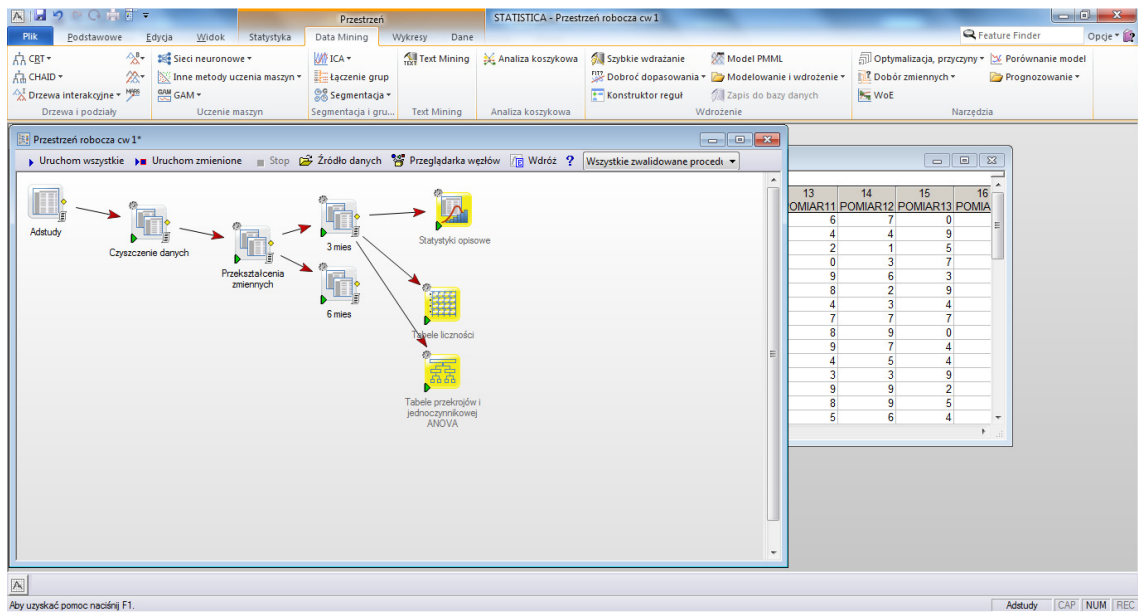
4. A w drugim sortuj według Średnia2
5. Uruchom i sprawdź wyniki

Teraz możemy już wykorzystać posortowane arkusze danych do kontaktów z odpowiednimi osobami, "najbardziej zadowolone" osoby są na górze.

Podłączanie węzłów analiz statystycznych

Dołączymy teraz do tych plików danych pewne statystyki opisowe i wykresy. Przypuśćmy, że interesuje nas ilość zakupów w ostatnich 3 oraz 6 miesiącach dla wszystkich obserwacji z pliku danych (czyli wszystkich kierowników zaopatrzenia), oraz średnie, w podziale względem zmiennej *REKLAMA* (podział na *Coke* oraz *Pepsi*).

1. Wstaw węzły: Statystyki -> Podstawowe -> Statystyki podstawowe -> Statystyki opisowe, Tabele liczości, Tabele przekrojów i jednoczynnikowej ANOVA
2. Skonfiguruj węzły:
 - a. Statystyki opisowe i tabele liczości wybierz Średnia1, Średnia2, REKLAMA
 - b. Tabele przekrojów i jednoczynnikowej ANOVA wybierz jako zmienne zależne: Średnia 1, Średnia2, a zmienne grupujące: REKLAMA



Przykładowo, reklama *Pepsi* na przestrzeni ostatnich 3 miesięcy jest bardziej skuteczna niż na przestrzeni 6 miesięcy. Oprócz tego podane są tabele licznosci itp.