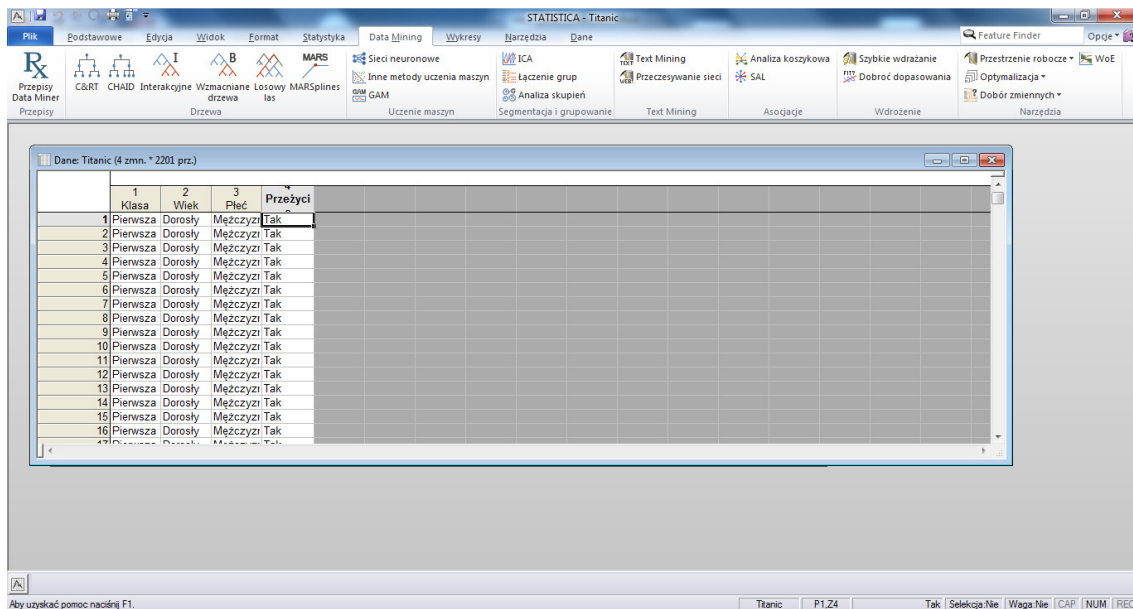


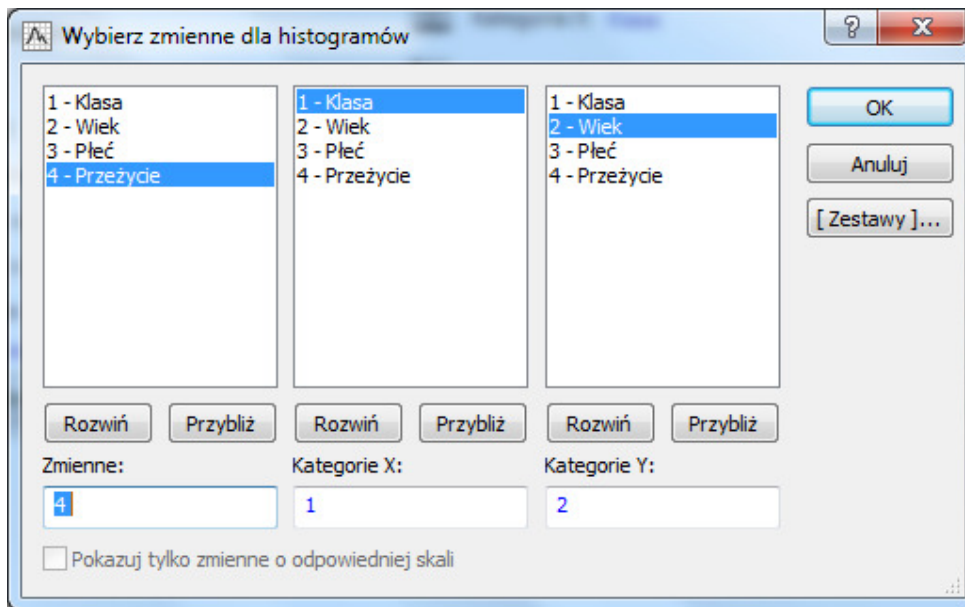
Ćwiczenie 1: Graficzny data mining

Celem przykładu jest przedstawienie sposobu wizualizacji danych w *STATISTICA Data Miner* (jak stosunkowo proste wykresy ujawniają wiedzę o związkach między zmiennymi). System *STATISTICA Data Miner* zawiera wiele wykresów, które można wykorzystywać razem z innymi narzędziami programu i łączyć w ten sposób "graficzny data mining" z technikami obliczeniowymi.

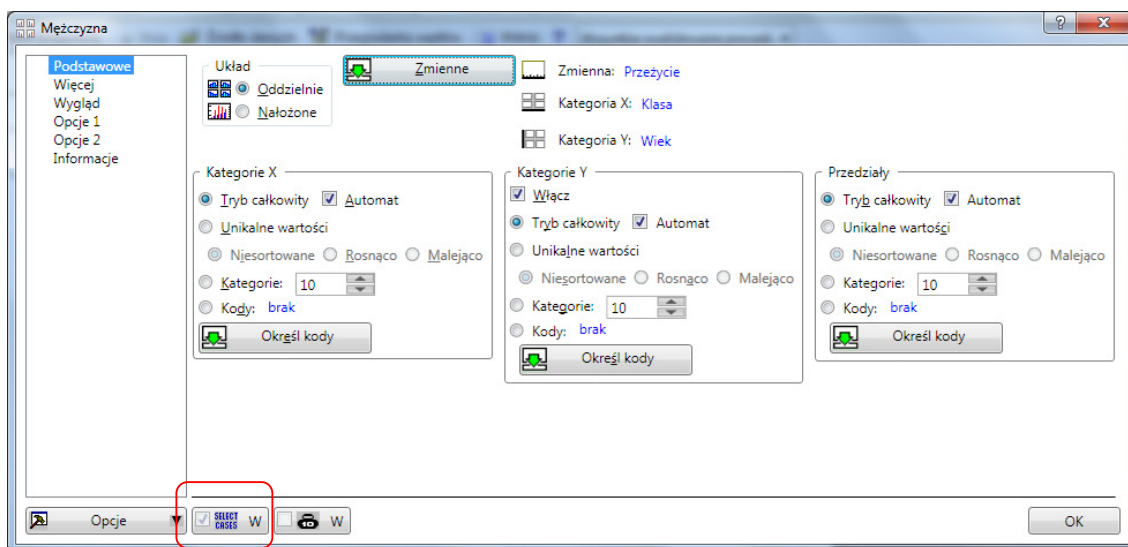
W przykładzie wykorzystamy dane z pliku *Titanic.sta*. W pliku zapisano informacje o płci, wieku i rodzaju zaokrętowania (np. *Pierwsza klasa*, *Obsługa*) osób podróżujących statkiem Titanic oraz czy przeżyły katastrofę, czy nie. Celem graficznego data mining będzie zbadanie proporcji przeżywających pasażerów z różnych grup.

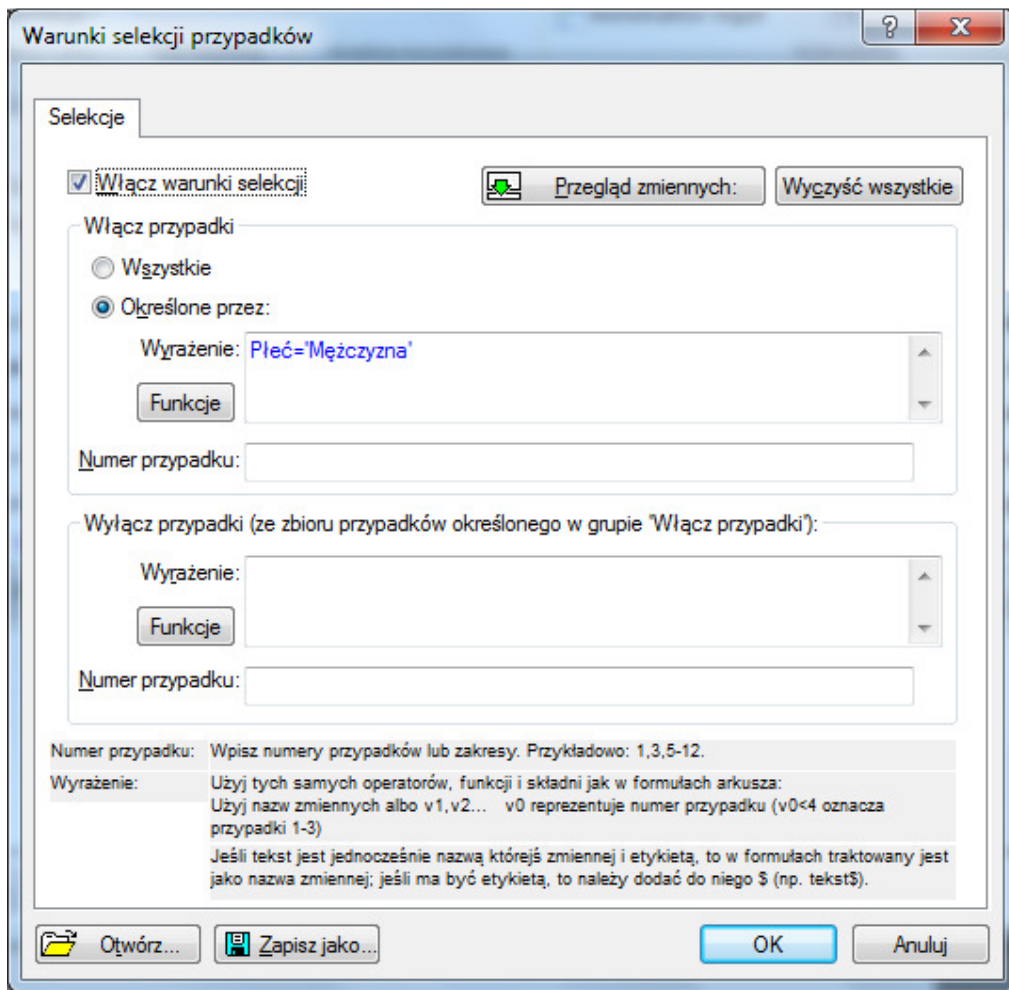


1. Otwórz przestrzeń roboczą
2. Wstaw źródło danych
3. Otwórz przeglądarkę węzłów
4. Wstaw Wykresy -> 2W -> Histogramy 2W
5. Wybierz wszystkie zmienne
6. Wstaw 2 węzły Wykresy -> Skategoryzowane -> Skategoryzowane histogramy
7. W pierwszym z nich wybierz zmienne:



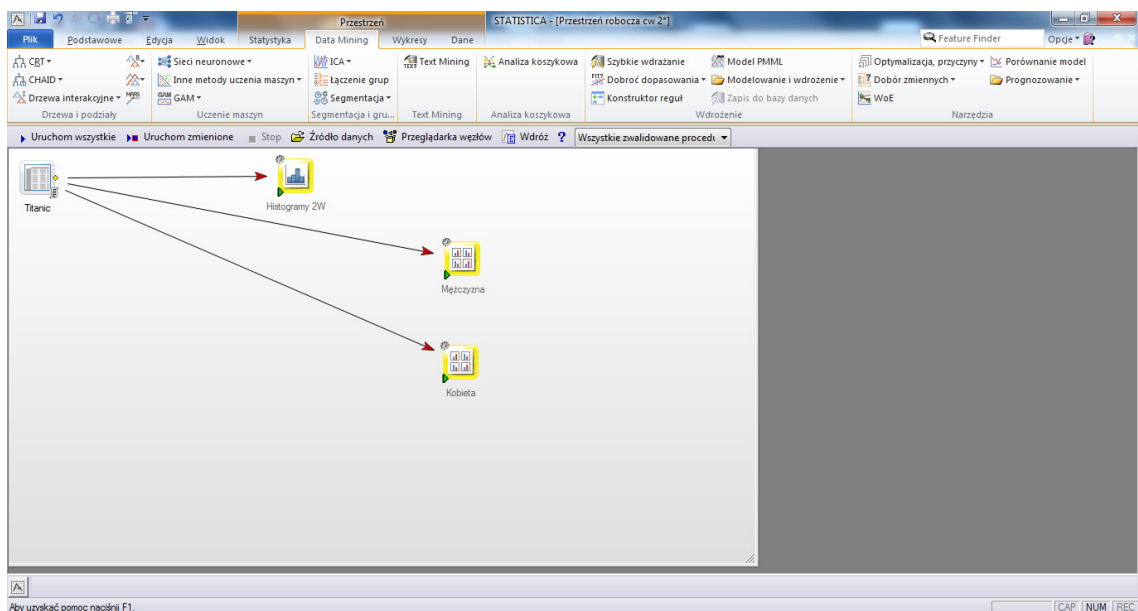
a następnie wybierz tylko mężczyzna (Select cases):



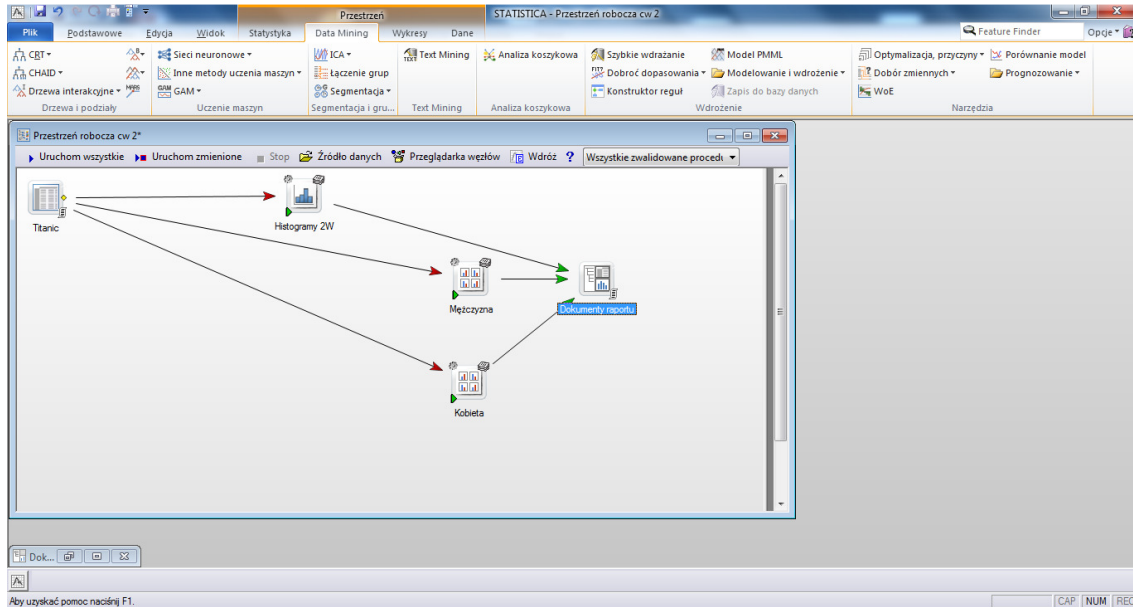


Zmień nazwę węzła na „Mężczyzna”

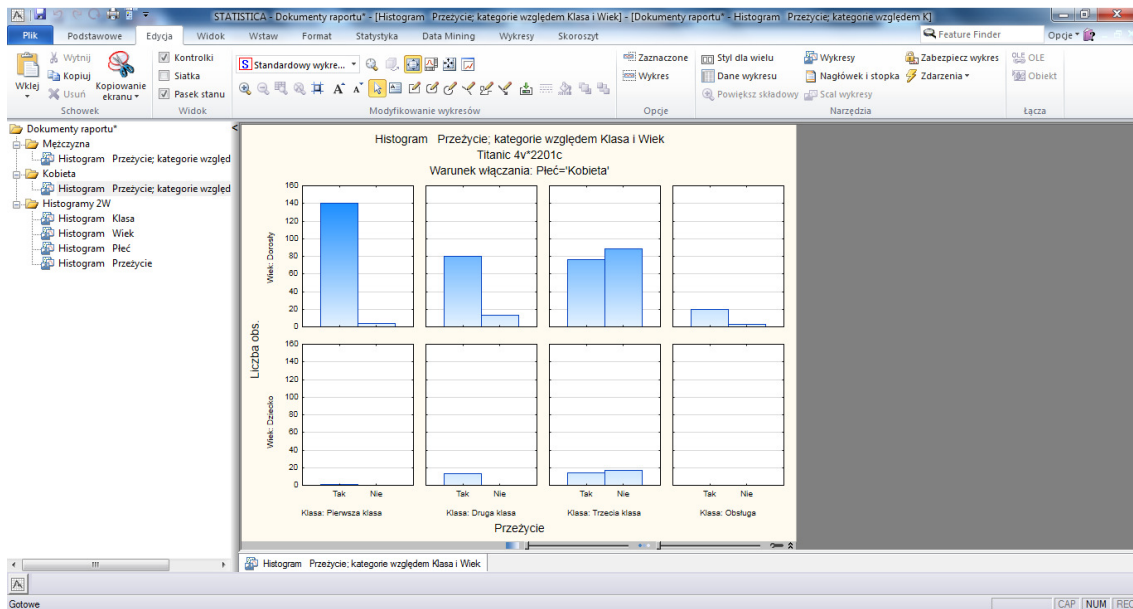
8. Podobnie w drugim węźle wybierz kobiety i zmień nazwę węzła na „Kobiety”:



9. Uruchom wszystkie węzły



Przeglądanie wyników. Po dwukrotnym kliknięciu możemy przejrzeć skoroszyty z wynikami. Zapewne najbardziej odkrywczym wykresem jest histogram skategoryzowany dla kobiet.



Rzuca się w oczy wyraźnie większy udział kobiet, które przeżyły katastrofę i podróżowały w pierwszej klasie, w porównaniu do kobiet które podróżowały w drugiej, a szczególnie trzeciej klasie. Pasażerki z trzeciej klasy miały wyraźnie najniższą szansę przeżycia: ponad 50% z nich zginęło, tymczasem w przypadku kobiet podróżujących w pierwszej i drugiej klasie znacząca większość wyszła cało z wypadku. Niewiele kobiet wchodziło w skład załogi statku

i większość z nich przeżyła katastrofę. W przypadku mężczyzn sytuacja jest diametralnie różna: 3/4 męskiej załogi zginęło w katastrofie. Wpływ klasy, w której podróżował pasażer nie ma aż tak silnego wpływu na procent mężczyzn przeżywających katastrofę. We wszystkich klasach większość mężczyzn zginęła, zginęła także zdecydowana większość męskiej załogi. Wyraźnie widać, że największe szanse przeżycia miały kobiety podróżujące pierwszą klasą (co może być świadectwem rycerskości tam czasów), a najmniejsze szanse pasażerowie trzeciej klasy i załoga.