



## **STATISTICA W ADMINISTRACJI PUBLICZNEJ**

*Piotr Wójtowicz, StatSoft Polska Sp. z o.o.*

Administracja publiczna to jeden z obszarów, gdzie odpowiednio zastosowane metody analizy danych mogą przynieść znaczne korzyści. Dlatego narzędzia *STATISTICA* są tak chętnie wykorzystywane przez pracowników administracji publicznej różnych szczebli: od ministerstw, po niewielkie jednostki samorządu terytorialnego. W tym artykule skoncentrujemy się na wskazaniu najważniejszych obszarów, w jakich analiza danych i *STATISTICA* mogą przynieść korzyści w administracji publicznej, oraz na tym, które aspekty oferty StatSoft Polska będą w tych zastosowaniach najefektywniejsze.

### **Obszary analizy danych w administracji**

Przedstawiciele administracji publicznej coraz lepiej zdają sobie sprawę z rosnącej roli informacji w zarządzaniu instytucjami sektora publicznego. Zastosowanie odpowiednich narzędzi analitycznych umożliwia, z wykorzystaniem zebranych informacji, określenie właściwych celów działania władz publicznych.

StatSoft Polska nie poprzestaje na oferowaniu swoim klientom narzędzi analitycznych z rodziny *STATISTICA*. Bardzo ważnym elementem naszej działalności jest prowadzenie szkoleń, konsulting w zakresie analizy danych oraz budowa kompletnych systemów analitycznych. Korzystanie nawet z zaawansowanego systemu analitycznego nie musi być trudne - pozostaje tylko odpowiedzieć na pytanie, gdzie analiza danych może przynieść wymierne korzyści.

**Sprawozdawczość.** Narzędzia analizy danych zawarte w *STATISTICA* pomagają zarówno w raportowaniu rutynowych działań organizacji, jak również w monitorowaniu stanu i realizacji projektów.

**Analizy przekrojowe.** Dzięki wykorzystaniu narzędzi analitycznych *STATISTICA* można przeprowadzać analizy przekrojowe, analizy w grupach, porównywać różne grupy. Dzięki specjalnie zaprojektowanemu interfejsowi wykonywanie tego typu analiz nie wymaga rozległej wiedzy na temat statystyki.

**Prognozowanie.** Dzięki modelom prognostycznym nie tylko uzyskamy odpowiedź na pytania, jaki będzie przewidywany poziom różnych parametrów i wskaźników w przyszłości, ale także jakie czynniki wpływają na te wartości. Trafnie postawiona



prognoza może więc być podstawą lepszych decyzji i planów, ale także daje możliwość lepszego przygotowania się do spodziewanych zmian, jakie zajądą w przyszłości.

**Monitorowanie satysfakcji.** W przypadku organizacji sektora publicznego (i innych organizacji *non profit*) satysfakcja „klienta” stanowi wartość samą w sobie i jeden z głównych celów działania. Także w tym obszarze narzędzia analizy danych pozwalające analizować i monitorować poziom satysfakcji są niezbędne.

**Zarządzanie zasobami i optymalizacja kosztów.** Analiza danych pozwala na określenie działań, które przyniosą najlepsze efekty, a także na ocenę wyników podejmowanych działań, na przykład określenie, czy środki, którymi dysponowała dana jednostka, wydawane były w sposób racjonalny i efektywny.

**Wizualizacja danych.** Jeden wykres potrafi przekazać więcej niż dziesiątki tabel. *STATISTICA* od lat słynie z wysokiej jakości wykresów, które pozwalają na przygotowanie nie tylko merytorycznie zaawansowanych, ale także efektownych raportów. Dodatkowo możliwość prezentacji zależności znalezionych w danych za pomocą wykresów na mapach, są szczególnie przydatne w administracji publicznej.

**Badania ankietowe i badania opinii.** Instytucje publiczne często prowadzą badania ankietowe, w szczególności badania opinii publicznej pozwalające uzyskać bogatą i pewną wiedzę o preferencjach i potrzebach obywateli. Zakres tych badań obejmuje m.in.: badania społeczne, badania ekonomiczne, badania opinii i zachowań społecznych, badania opinii pracowniczych, badania polityczne. Narzędzia *STATISTICA* pozwalają na efektywne przeprowadzenie takich badań oraz szybką i wygodną analizę wyników.

**Upowszechnienie informacji.** W dobie powszechnego dostępu do Internetu niezwykle istotna jest kwestia upowszechnienia informacji gromadzonych przez instytucje administracji publicznej. Większość zgromadzonych zasobów danych jest niewykorzystana trafiając najczęściej do zasobów archiwalnych. Dzięki zastosowaniu narzędzi analitycznych można wydobyć te informacje, które są istotne.

**Wykrywanie nadużyć.** Dzięki analizie danych można wykryć elementy nietypowe, które mogą wskazywać na potencjalne nadużycia. Poprzez identyfikację grup ryzyka (np. grupy niskiego, średniego i wysokiego ryzyka) wykrywanie nadużyć i innych nielegalnych operacji może skupić się przede wszystkim na najbardziej zagrożonych przypadkach.

**Zarządzanie.** Kierownictwo, poza raportem opisującym aktualną sytuację, potrzebuje również dodatkowych informacji, które pozwoliłyby lepiej zrozumieć organizację, a tym samym sprawnie nią zarządzać. Taką wiedzę można uzyskać dzięki wykorzystaniu metod statystycznych do danych, które są dostępne w organizacji.

## Opinie klientów

Od wielu lat *STATISTICA* i StatSoft cieszą się dużym uznaniem u klientów, szczególnie administracji publicznej. Poniżej przedstawione są wybrane opinie naszych klientów:



### **Ministerstwo Finansów**

*Wynikiem przeprowadzonego projektu badawczego przez pracowników firmy StatSoft było dostarczenie raportu z przeprowadzonej analizy zawierającego przejrzystą interpretację wyników analizy oraz niezbędne tabele i wykresy ilustrujące zależności pomiędzy analizowanymi zmiennymi. Rezultaty wyników zawarte w raporcie oceniamy wysoko z uwagi na profesjonalne podejście, dużą wiedzę z zakresu metod statystycznych i właściwe wykorzystanie przekazanej przez nas wiedzy merytorycznej a także duże możliwości zastosowanych narzędzi oprogramowania analitycznego.*

*Jacek Krzyślak  
Dyrektor Departamentu Polityki Finansowej, Analiz i Statystyki  
Ministerstwo Finansów*

### **Wojewódzki Urząd Pracy w Gdańsku**

*Oprogramowanie to wyraźnie przewyższa funkcjonalnością inne aplikacje. Wielką wagę dla użytkownika ma intuicyjność, łatwość obsługi oraz wygoda użycia programu. Obszerny wybór narzędzi analitycznych w połączeniu z szeroką dokumentacją (zarówno w formie elektronicznej, jak i książkowej) pozwalają użytkownikowi szybko i sprawnie przeprowadzić potrzebne analizy. Szczególnie przydatny jest dla nas unikatowy na rynku moduł STATISTICA Mapy, umożliwiający wizualizację danych statystycznych dotyczących poszczególnych jednostek administracyjnych, gdyż przygotowywane w Urzędzie materiały opierają się w dużym stopniu na graficznej prezentacji danych.*

*Ewa Jurkowska  
Dyrektor  
Wojewódzki Urząd Pracy w Gdańsku*

### **Centralny Instytut Ochrony Pracy**

*Analiza, którą otrzymaliśmy została przeprowadzona w sposób wnikliwy i kompetentny. Jesteśmy szczególnie wdzięczni za przygotowanie profesjonalnego raportu z przeprowadzonej analizy zawierającego przejrzystą interpretację wyników analizy oraz niezbędne tabele i wykresy ilustrujące zależności między analizowanymi zmiennymi.*

*Anna Marszałek  
Centralny Instytut Ochrony Pracy  
-Państwowy Instytut Badawczy*

### **EUROTRANSFER Registry**

*EUROTRANSFER Registry to ogólnoeuropejski rejestr pacjentów z zawałem mięśnia sercowego. Prace związane z analizą danych badawczych zostały powierzone specjalistom z firmy StatSoft Polska. Projekt obejmował całość procesu analizy.*

*Doceniam stopień zaangażowania pracowników StatSoft w całość projektu. To ich szeroka wiedza i doświadczenie, ale także elastyczność, były gwarancją powodzenia projektu.*



*Uważam, że tak poważne przedsięwzięcie, jakim jest analiza danych badawczych, powinno być realizowane przez profesjonalistów – polecam konsultantów firmy StatSoft Polska.*

*Lukasz Partyka  
Zbigniew Siudak  
Krakow Cardiovascular Research Institute*

### ***Krajowy Depozyt Papierów Wartościowych***

*Program STATISTICA, a w szczególności STATISTICA Data Miner zautomatyzował procedury sporządzania prognoz finansowych w naszej firmie, powodując jednocześnie, że ich wykonywanie stało się bardziej przejrzyste i mniej pracochłonne. Dzięki możliwości modelowania tak dużej ilości szeregów jednocześnie nasze prognozy są dokładniejsze a ich prezentacja graficzna jest przyjemna, prosta i efektowna.*

*Jolanta Semeniuk-Bijak  
Dyrektor Działu Analiz i Studiów  
Krajowy Depozyt Papierów Wartościowych S.A.*

### ***Narodowy Bank Polski***

*Departament Kadr Narodowego Banku Polskiego potwierdza, iż od 2002 roku współpracuje przy realizacji szkoleń z firmą StatSoft Polska Sp. z o.o. z Krakowa. Do chwili obecnej pracownicy Centrali NBP skorzystali z wielu szkoleń oferowanych przez firmę StatSoft realizowanych w siedzibie Firmy, jak również szkoleń realizowanych w NBP zgodnie z programami dostosowanymi do potrzeb i specyfiki banku centralnego.*

*Firma StatSoft angażuje uznanych wykładowców o znaczącym dorobku naukowym, bardzo dużym doświadczeniu zawodowym i praktycznym, którzy wykorzystując oprogramowanie STATISTICA w sposób jasny i komunikatywny wyjaśniali specjalistyczne zagadnienia z zakresu ekonometrii, matematyki finansowej, statystyki.*

*Firma StatSoft zapewniła wysoki poziom merytoryczny i elastyczność w zakresie doboru treści programowych szkoleń dedykowanych dla NBP. Przeprowadzone zajęcia zostały bardzo wysoko ocenione i w pełni satysfakcjonują naszych pracowników.*

*Jerzy Kerszke  
Dyrektor Departamentu Kadr  
Narodowy Bank Polski*

## **Rozwiązania StatSoft**

*STATISTICA to rodzina aplikacji zapewniająca kompletne rozwiązania w zakresie analizy danych – najlepsze narzędzia analityczne, systemy korporacyjne z elementami pracy grupowej i kontroli dostępu oraz możliwością zdalnej pracy przez Internet, mechanizmy dostępu do baz i hurtowni danych oraz systemy rozpowszechniania informacji. Wybierając poszczególne produkty z rodziny STATISTICA, można zbudować rozwiązanie dostosowane do aktualnych potrzeb użytkowników.*



*STATISTICA* spełnia wszystkie wymagania jakie jednostki administracji publicznej stawiają systemom analitycznym:

- ◆ dostarcza proste w obsłudze, intuicyjne środowisko pracy;
- ◆ udostępnia złożone narzędzia wizualizacji danych;
- ◆ zapewnia otwartość i elastyczność umożliwiającą wymianę danych pomiędzy różnymi aplikacjami, w tym MS Excel;
- ◆ jest zgodna z normami stosowanymi w administracji publicznej, w tym z rozwiązaniami z zakresu systemów zarządzania jakością opartymi na serii norm ISO 9000.

### ***Ogólne cechy programów z rodziny STATISTICA***

*STATISTICA* to rodzina programów wspomagających analizy danych wszelkiego rodzaju, od obliczania podstawowych charakterystyk i wizualizacji danych, poprzez modelowanie klasycznymi metodami statystycznymi, po wykorzystanie nowoczesnych technik optymalizujących pracę z bardzo dużymi zbiorami danych.

System analizy danych *STATISTICA* można zbudować zestawiając takie narzędzie analityczne jak jest potrzebne do konkretnego przypadku. Od zestawu do prostych zastosowań *STATISTICA Pakiet Podstawowy* po obszerne i zaawansowane narzędzie data mining *STATISTICA Data Miner*. *STATISTICA* to także możliwość zdalnej pracy przez Internet, współpracy z intranetem oraz elektronicznym obiegiem dokumentów.

### ***Środowisko pracy***

*STATISTICA* znana jest z prostego w użyciu, intuicyjnego, graficznego środowiska pracy. Można je z łatwością dostosować do własnych potrzeb i preferencji. Możliwe jest również definiowanie całkowicie nowego środowiska, w tym także interfejsów bazujących na przeglądarce internetowej. Wysokiej jakości narzędzia wizualizacji wyników analiz umożliwiają czytelną prezentację wyników i danych. Na każdym etapie analizy dostępny jest „pod ręką” zestaw charakterystycznych dla danej metody wykresów analitycznych, które umożliwiają przedstawienie w przejrzysty sposób wyników analizy. Oprócz wykresów wyników dostępny jest również bogaty zestaw wykresów prezentujących zgromadzone dane.

### ***Architektura***

Podstawowa architektura rozwiązania, niezależnie od wybranego wariantu w zakresie dostępnych technik analitycznych, to wersja sieciowa *concurrent*. Zapewnia ona efektywne wykorzystanie licencji – limitowana jest liczba użytkowników w danym momencie, a nie całkowita liczba użytkowników (oprogramowanie można zainstalować na dowolnej liczbie stacji roboczych). *STATISTICA* w wersji *concurrent* jest programem pozwalającym korzystać jednocześnie z systemu wielu użytkownikom, pracującym w lokalnej sieci Windows.

Dla instytucji posiadających mniejsze wymagania w zakresie korzystania z rozwiązania alternatywnie proponujemy wersję jednostanowiskową. Można ją zainstalować i wykorzystywać na jednym komputerze.



## ***STATISTICA Enterprise***

Jest to kompleksowy system analityczno-raportujący. *STATISTICA Enterprise* umożliwia pracę zespołową – w praktyce może to wyglądać w ten sposób, że pracownicy zajmujący się analizą danych przygotowują raporty, które następnie mogą być w prosty sposób udostępniane zwierzchnikom. *STATISTICA Enterprise* to wieloużytkownikowa wersja systemu umożliwiająca efektywną pracę grupową, automatyczne monitorowanie danych, automatyczne rozpowszechnianie danych, raportów oraz wyników analiz, przy zachowaniu bezpieczeństwa informacji. Bliższy opis znajduje się w dalszej części artykułu.

## ***Dane***

*STATISTICA* w każdej wersji zawiera *STATISTICA Query*, narzędzie które umożliwia łatwy dostęp do danych gromadzonych w różnego typu bazach danych (np. Sybase, Oracle, MS SQL Server itp.) za pomocą standardu OLE DB lub ODBC. *STATISTICA* pozwala szybko i w prosty sposób tworzyć rozbudowane zapytania bez znajomości języka SQL.

*STATISTICA* pozwala także na pracę z plikami danych z rozmaitych aplikacji, np. MS Excel, pliki tekstowe, .xml i wiele innych.

*STATISTICA* jest zoptymalizowana pod kątem pracy z dużymi plikami danych. Dane przechowywane są w arkuszach o praktycznie nieograniczonej pojemności. Nie ma również ograniczeń co do zawartości poszczególnych komórek arkusza, można do nich dołączać pliki dźwiękowe, wideo, wykresy, animacje oraz inne obiekty zgodne z ActiveX. Dostępny opcjonalnie mechanizm IDP pozwala także na pracę z danymi (zgromadzonymi np. w bazie czy hurtowni danych) bez konieczności tworzenia ich lokalnej kopii.

## ***Automatyzacja***

Ponadto aplikacja ma wbudowane środowisko programistyczne *STATISTICA Visual Basic* (zgodne ze standardowym MS Visual Basic), które w pełni wykorzystuje zalety modelu obiektowego *STATISTICA* i umożliwia programowy dostęp do praktycznie każdego elementu i każdej funkcji programu.

## ***Otwartość***

Otwarta architektura pozwala na integrację *STATISTICA* z praktycznie dowolnym systemem gromadzenia danych, StatSoft oferuje budowę dedykowanych systemów analitycznych i ich integrację z istniejącymi systemami.

## ***Prezentacja wyników***

W trakcie przeprowadzania analiz w *STATISTICA* na wyjściu otrzymujemy wyniki w postaci multimedialnych tabel (arkuszy) i wykresów.

*STATISTICA* od lat jest bezkonkurencyjna wśród narzędzi do statystycznej analizy danych i data mining pod względem narzędzi wizualizacji danych i tworzenia wykresów. *STATISTICA* oferuje setki typów dwuwymiarowych i trójwymiarowych wykresów oraz



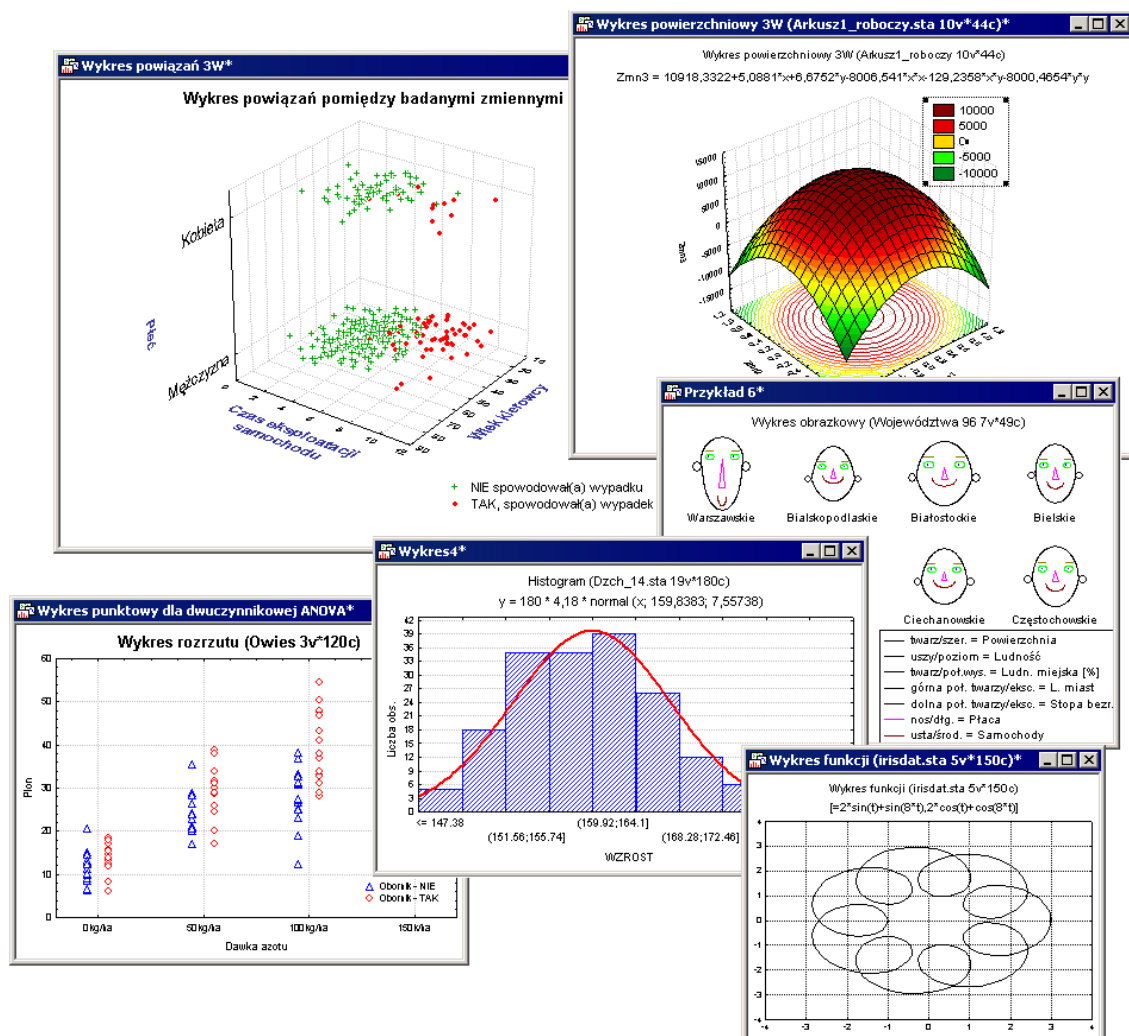
wygodne narzędzia do projektowania nowych typów własnych wykresów i dodawania ich do menu lub pasków narzędzi. Oprócz ogólnych typów wykresów istnieje wiele specjalistycznych wykresów, zintegrowanych ze specjalistycznymi procedurami analitycznymi i dostępnych w odpowiednich oknach wyboru wyników lub w podręcznych menu. Praktycznie każdy szczegół wyglądu wykresu (a są to setki cech) może być ustawiany przez użytkownika. Wszystkie wykresy utworzone na podstawie arkusza danych mogą być automatycznie odświeżane przy każdej modyfikacji danych lub na życzenie użytkownika.

Wyniki w postaci tabel prezentowane są w postaci multimedialnych arkuszy w stylu MS Excel.

Raport *STATISTICA* przypomina dokument edytora tekstu, a poszczególne obiekty (np. wykresy, arkusze, arkusz czy wykres MS Excel) umieszczane są w nim kolejno, jeden za drugim. Wszystkie raporty mogą być zapisywane nie tylko we własnym formacie *STATISTICA*, ale także w postaci plików DOC, RTF, HTML i PDF.

### Najwyższej jakości grafika

Dostępne w każdym produkcie z rodziny *STATISTICA* wysokiej jakości narzędzia wizualizacji wyników analiz umożliwiają czytelną prezentację wyników i danych.





Na każdym etapie analizy mamy „pod ręką” zestaw charakterystycznych dla danej metody wykresów analitycznych, które umożliwiają przedstawienie w przejrzysty sposób wyników danej analizy. Oprócz wykresów wynikowych mamy również do wyboru bogaty zestaw wykresów prezentujących zgromadzone dane.

### ***Skalowalność***

Rozwiązania z rodziny *STATISTICA* są skalowalne pod każdym względem:

- ◆ liczby użytkowników – od wersji jednostanowiskowych po korporacyjne systemy wieloużytkownikowe z kontrolą dostępu i pracą przez Internet,
- ◆ wielkości przetwarzanych danych – *STATISTICA* jest zoptymalizowana pod kątem szybkości wykonywanych analiz, specjalna technologia IDP pozwala na analizę ekstremalnie dużych zbiorów danych bez konieczności tworzenia lokalnej kopii całego analizowanego zbioru,
- ◆ zakresu procedur analitycznych – w każdym momencie można rozszerzyć licencję o dodatkowe metody analityczne, cały czas pracując w tym samym środowisku, bez konieczności zmiany interfejsu użytkownika i związanych z nim przyzwyczajzeń.

### ***Przetwarzanie równoległe i rozproszone***

*STATISTICA* w wersji klient-serwer obsługuje przetwarzanie równoległe i rozproszone po stronie serwera. Technologia w pełni równoległego przetwarzania rozdziela zadania nie tylko pomiędzy wiele procesorów serwera, ale również na wiele komputerów, co radykalnie zwiększa wydajność w projektach o wysokich wymaganiach obliczeniowych, czy takich, w których wykorzystywane są wielkie zbiory danych.

### ***Programowalność i otwarta architektura***

Każdy z programów z rodziny *STATISTICA* ma wbudowane środowisko programistyczne *STATISTICA Visual Basic* (zgodne ze standardowym MS Visual Basic), które w pełni wykorzystuje zalety modelu obiektowego *STATISTICA* i umożliwia programowy dostęp do praktycznie każdego elementu i każdej funkcji programu. *STATISTICA Visual Basic* rozszerza standardową składnię języka Visual Basic o ponad 11 tysięcy nowych funkcji (w wersji 8 *STATISTICA* dodane zostały kolejne procedury), co daje bardzo rozbudowane i bogate środowisko programistyczne. Umożliwia ono tworzenie specjalizowanych systemów analitycznych, które łączą potężne możliwości zoptymalizowanych procedur *STATISTICA* z różnymi rozszerzeniami i własnym interfejsem. Środowisko *STATISTICA Visual Basic* oferuje programiście elastyczny edytor i debugger, edytor okien dialogowych, a także przeglądarki dostępnych funkcji i obiektów.

Programy *STATISTICA Visual Basic* można uruchamiać nie tylko w *STATISTICA*, lecz (dzięki zgodności ze standardami) także w innych środowiskach bazujących na konwencjach Microsoft Visual Basic (np. MS Excel, MS Word lub w samodzielnych instalacjach Visual Basic). Należy zauważyć, że ta bogata biblioteka funkcji *STATISTICA* jest dostępna nie tylko dla aplikacji w języku Visual Basic (wbudowanego w program lub też innego),

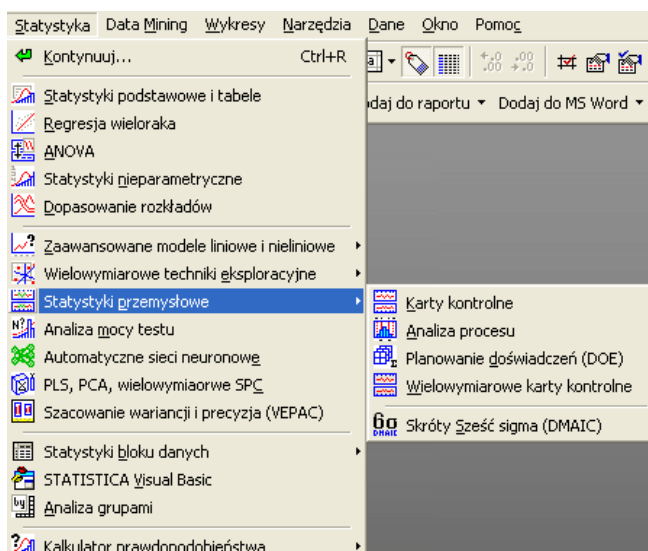


lecz również z dowolnego innego zgodnego środowiska programistycznego (np. C/C++, Java, Delphi).

## Narzędzia analityczne

Każda z aplikacji analitycznych może pracować zarówno jako wersja jednostanowiskowa (na pojedynczym komputerze), w sieci (na wielu komputerach użytkowników), jak i może być elementem korporacyjnego systemu analizy danych.

Zakres metod analitycznych oferowanych przez *STATISTICA* jest niezwykle bogaty. Metody analityczne zostały podzielone na różne programy, z których można zbudować odpowiedni zestaw analityczny. Poszczególne programy z rodziny *STATISTICA* można łączyć w jednym pakiecie, który może z czasem być rozszerzany o dodatkowe moduły. Dzięki temu możemy mieć dostęp do bardzo wielu różnych metod analitycznych zawartych w jednej aplikacji *STATISTICA*.

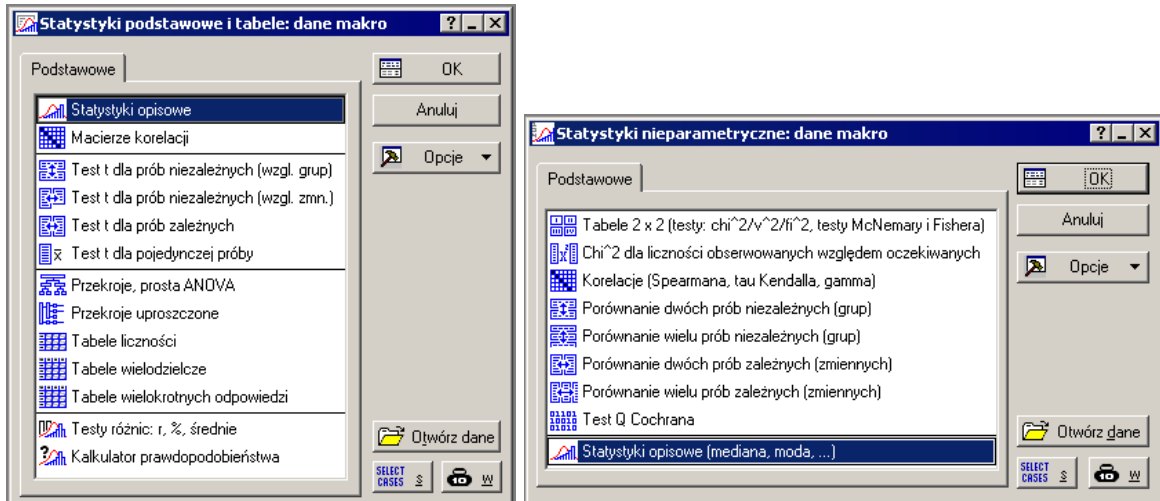


Poniżej znajdują się opisy wybranych aplikacji *STATISTICA*, najczęściej stosowanych w administracji publicznej. Opisy wszystkich programów z rodziny *STATISTICA* dostępne są na stronie internetowej [www.StatSoft.pl](http://www.StatSoft.pl).



## *STATISTICA* Pakiet Podstawowy PL

*STATISTICA* Pakiet Podstawowy to obszerny zestaw najpopularniejszych metod statystycznych, obejmujący: *Statystyki opisowe*, *Tabele liczebności*, *Przekroje*, *Tabele wielodzielcze*, *Korelacje*, *Testy parametryczne*, *ANOVA*, *Analiza regresji*, *Dopasowanie rozkładów*, *Testy nieparametryczne* i *Kalkulator prawdopodobieństwa*. Zawiera pełne środowisko *STATISTICA*: modyfikowalny interfejs, elastyczne narzędzia zarządzania wynikami i raportowania, interakcyjne narzędzia do tworzenia zapytań do baz danych (*STATISTICA Query*), wszechstronne narzędzia importu i eksportu danych, komplet narzędzi wizualizacji danych i wyników analiz (wysokiej jakości wykresów) oraz w pełni zintegrowany język programowania *STATISTICA Visual Basic*.

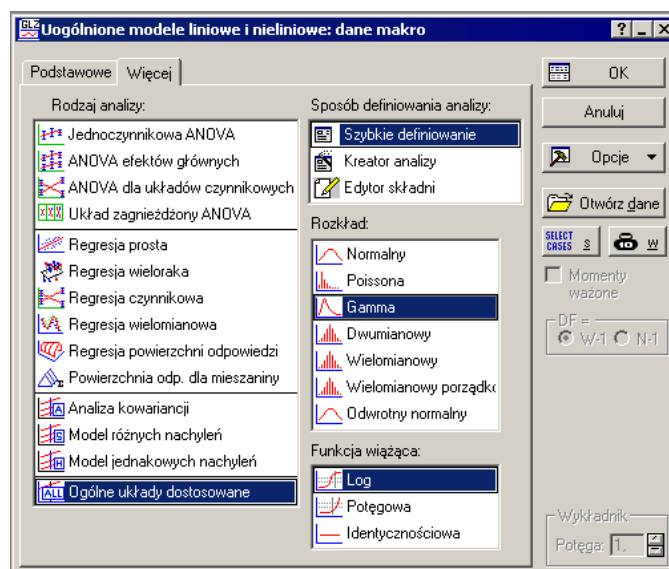


## STATISTICA PL

Zestaw *STATISTICA* zawiera wszystkie funkcjonalności programu *STATISTICA Pakiet Podstawowy*, a także moduły *Modele Zaawansowane*, *Analizy Wielowymiarowe* oraz *Analiza Mocy Testów*.

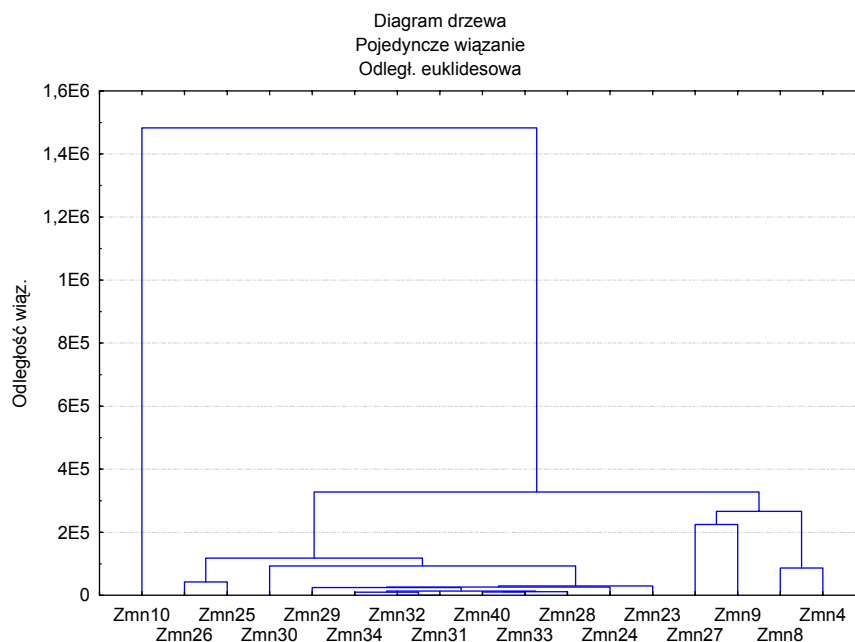


*Modele Zaawansowane* to szeroki wybór zaawansowanych narzędzi do budowy modeli liniowych i nieliniowych. Umożliwiają one wykorzystanie predyktorów ciągłych i jakościowych, uwzględnianie interakcji, tworzenie modeli hierarchicznych oraz automatyczny wybór modeli. Ponadto pozwalają obliczać komponenty wariancyjne, analizować dane o postaci szeregów czasowych oraz stosować wiele innych metod. Moduł obejmuje następujące grupy procedur: *Ogólne modele liniowe (GLM)*, *Uogólnione modele liniowe i nieliniowe*, *Ogólne modele regresji*, *Szeregi czasowe i prognozowanie*, *Modele cząstkowych najmniejszych kwadratów*, *Algorytm NIPALS (PCA/PLS)*, *Komponenty wariancyjne i model mieszany ANOVA/ANCOVA*, *Analiza czasu przeżycia i awaryjności*, *Estymacja nieliniowa*, *Linearyzowana regresja liniowa*, *Analiza log-liniowa tabel liczości*, *Modelowanie równań strukturalnych i analiza ścieżkowa (SEPATH)*.





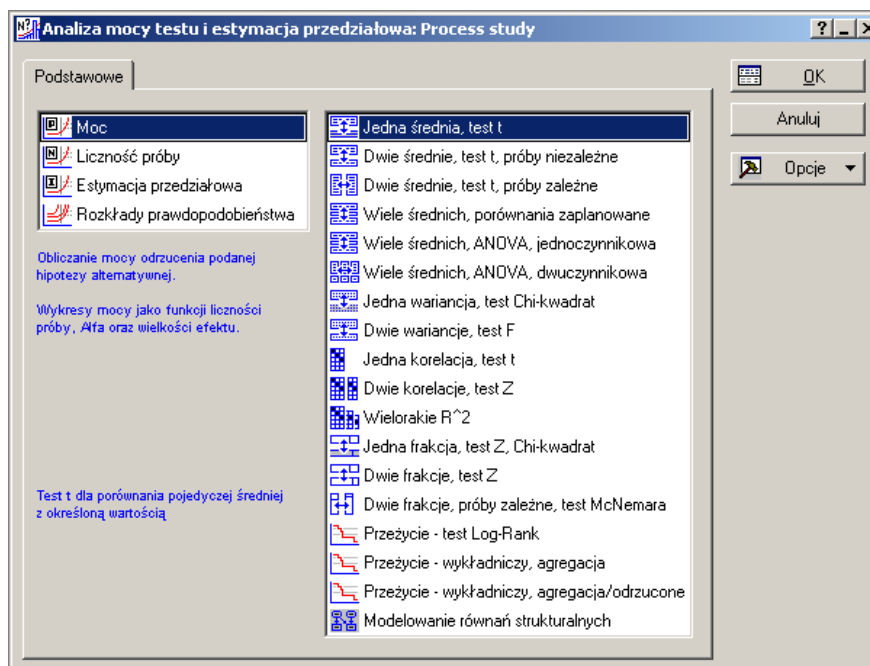
*Analizy Wielowymiarowe* to bogaty zestaw technik eksploracji danych różnego typu, począwszy od analizy skupień, a skończywszy na wybranych metodach drzew klasyfikacyjnych, wraz z obszernym zestawem interakcyjnych narzędzi wizualizacji, służących do badania współzależności i występowania regularnych układów danych. Moduł obejmuje następujące grupy procedur: *Analiza skupień*, *Analiza czynnikowa*, *Analiza składowych głównych i klasyfikacja*, *Analiza kanoniczna*, *Analiza rzetelności i pozycji*, *Drzewa klasyfikacyjne*, *Analiza korespondencji*, *Skalowanie wielowymiarowe*, *Analiza dyskryminacyjna*, *Ogólne modele analizy dyskryminacyjnej (GDA)*.



Przedstawiony powyżej wykres przedstawia przykładowy diagram drzewa będący wynikiem *Analizy skupień*.



*Analiza Mocy Testów* to specjalistyczne, a jednocześnie przyjazne dla użytkownika narzędzie do analizowania wszelkich aspektów mocy testów statystycznych i wyznaczania liczności próby. Moduł ten obejmuje obszerny zestaw narzędzi ułatwiających takie zaplanowanie badań, aby wielkość próby odpowiadała ich celowi, a także analizy związane z obliczaniem mocy statystycznej i wielkości próby. *STATISTICA Analiza Mocy Testów* wychodzi poza standardowe testy „braku różnic” i wykorzystuje także techniki bazujące na estymacji przedziałowej.



### **STATISTICA Automatyczne Sieci Neuronowe PL**

Program *STATISTICA Automatyczne Sieci Neuronowe* jest wydajną, a zarazem łatwą w obsłudze aplikacją wykorzystującą najnowszą technologię i algorytmy tworzenia oraz wdrażania modeli bazujących na sieciach neuronowych. Dzięki niej nawet użytkownicy niebędący specjalistami w zakresie sieci neuronowych mogą wykorzystywać te techniki sztucznej inteligencji do rozwiązywania rozmaitych problemów analitycznych. Na wstępie użytkownik wybiera rodzaj zagadnienia, jakie będzie analizowane. Można wybrać spośród następujących opcji:

- ◆ regresja (dla danych nietworzących szeregu czasowego),
- ◆ klasyfikacja (dla danych niesekwencyjnych),
- ◆ regresja szeregów czasowych,
- ◆ klasyfikacja szeregów czasowych,
- ◆ analiza skupień (z użyciem map Kohonena), czyli segmentacja.

Do najważniejszych funkcji programu *STATISTICA Automatyczne Sieci Neuronowe* należą m.in. zintegrowane narzędzia do wstępnego przetwarzania danych i wyników, automatyczny projektant sieci, algorytmy uczące zbudowane na bazie najnowszych rozwiązań, w wysokim stopniu zoptymalizowane (w tym algorytm gradientów sprzężonych i BFGS), możliwość wyboru funkcji aktywacji, funkcji błędu i złożoności sieci, możliwość konstruowania z sieci różnego typu zestawów sieci złożonych o praktycznie nieograniczonej wielkości, szeroki wybór narzędzi graficznych i statystycznych umożliwiających interakcyjną analizę eksploracyjną, pełna integracja z systemem *STATISTICA*.



The screenshot displays the STATISTICA software interface. The main window is titled "STATISTICA - [Dane: IrisSNN.sta (6 zmn. \* 150 prz.)]". The menu bar includes "Plik", "Edycja", "Widok", "Wstaw", "Format", "Statystyka", "Data Mining", "Wykresy", "Narzędzia", "Dane", "Okno", and "Pomoc". The toolbar contains various icons for file operations, editing, and analysis. The main workspace is divided into several panels:

- Analiza lub wdrożenie SANN: IrisSNN.sta**: A dialog box for configuring a new SANN analysis. It includes options for "Wdrożenie" and "Nowa analiza", a list of analysis types (Regresja, Klasyfikacja, Szeregi czasowe (regresja), Szeregi czasowe (klasyfikacja), Analiza skupień), and a table for network parameters.
- SANN - Wyniki: IrisSNN.sta**: A window showing the results of the SANN analysis. It contains a table of active neural networks with columns for ID sieci, Nazwa sieci, Jakość (ucz.), Jakość (t.), Algorytm, Błąd, and Aktywacja... The table lists five networks with their respective parameters and activation functions.
- Skoroszyt1\* - Wykres zysku - Wybrane przez model/Wsz...**: A window displaying a decision tree plot. The plot shows a red line representing the predicted output and a blue diagonal line representing the actual output. The x-axis is labeled "Procent" and the y-axis is labeled "Zysk".

The bottom status bar shows "Gotowe", "IrisSNN.sta", and a list of variables: "Selekcja:Nie", "Waga:Nie", "CAP", "NUM", "ZAR".

Wszystkie wyniki, wykresy, raporty itp. mogą być przetwarzane za pomocą narzędzi graficznych i analitycznych *STATISTICA* (np. w celu analizy reszt z wyników prognoz, tworzenia podsumowujących raportów itd.). Program jest w pełni zintegrowany z narzędziami *STATISTICA* do automatyzacji analiz. Opcjonalnie dostępny jest generator kodu w języku C.

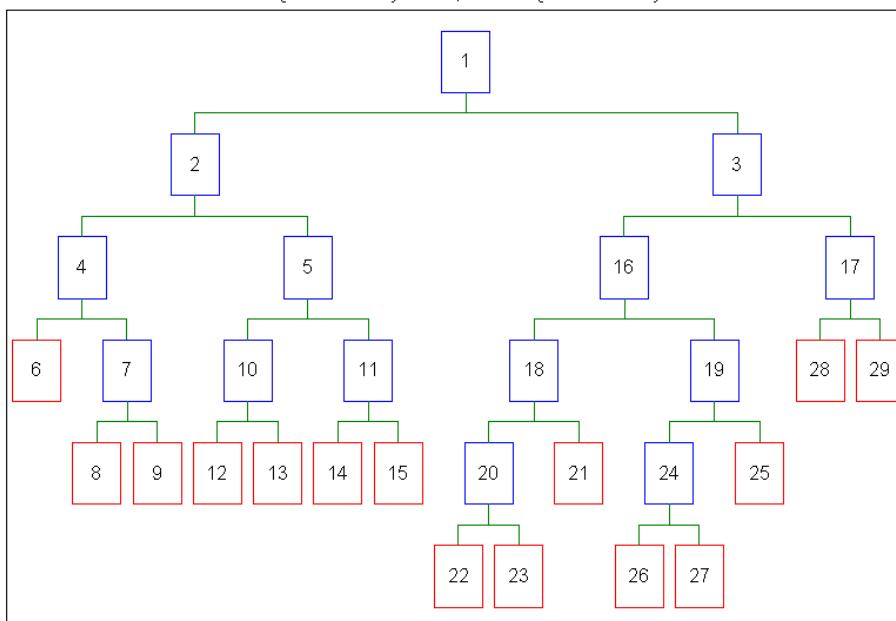
## *STATISTICA Trees PL*

Program *STATISTICA Trees* jest doskonały dla analityków chcących przeprowadzać analizy data mining przy wykorzystaniu modeli drzew decyzyjnych. W programie tym znajduje się część metod dostępnych także w *STATISTICA Data Miner*. Są to: *Ogólne modele drzew klasyfikacyjnych i regresyjnych (G-Trees)*, *Ogólne modele CHAID*, *drzewa interakcyjne* i *wzmacniane*. *STATISTICA Trees* umożliwia wykonywanie analiz z zastosowaniem algorytmu budowy drzew, które prowadzą do znalezienia zbioru logicznych warunków podziału typu „jeżeli, to”, pomagają jednoznacznie zaklasyfikować badane obiekty. Mechanizm eksploracji drzewa jest bardzo przydatny przy szczegółowej analizie wyników.



Układ drzewa 1 dla Zmn2

Liczba węzłów dzielonych: 14, liczba węzłów końcowych: 15

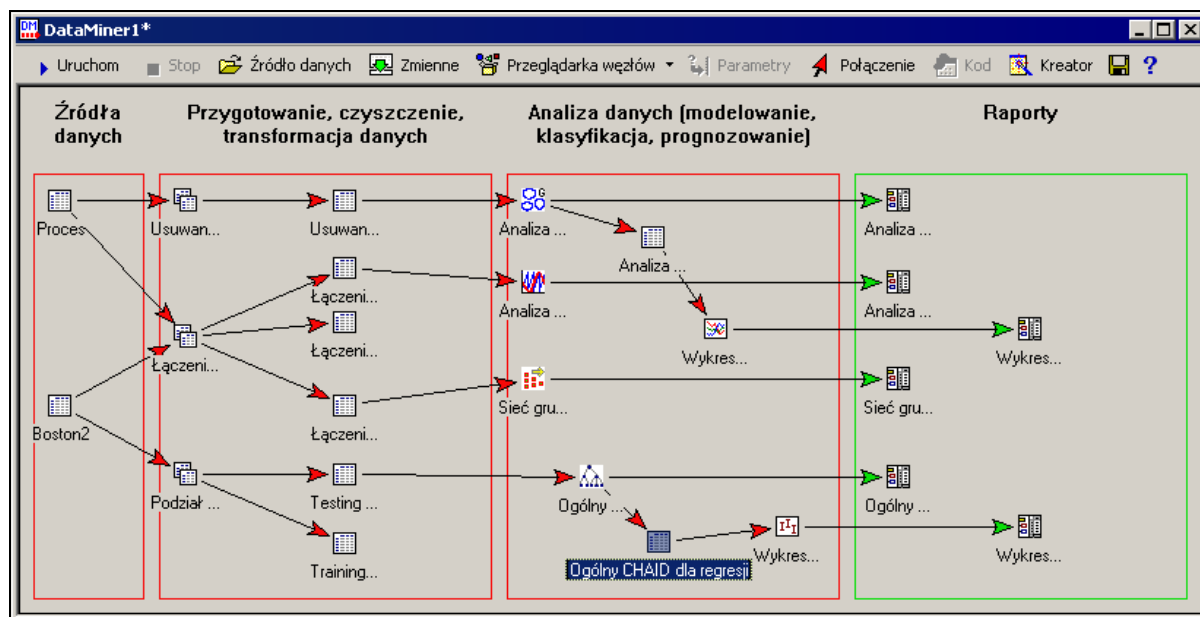


### **STATISTICA Data Miner PL**

*STATISTICA Data Miner* to kompletny zestaw narzędzi data mining, zaprojektowany tak, aby umożliwić łatwe i szybkie wykonywanie analiz danych i stosowanie wyników we wspomaganie podejmowania decyzji. *STATISTICA Data Miner* jest łatwym w użyciu, a jednocześnie bardzo wydajnym narzędziem do realizacji wszystkich etapów wydobycia z danych użytecznej wiedzy: począwszy od pobierania danych z baz i hurtowni danych, poprzez etap czyszczenia i transformacji danych, wykonywanie analiz danych, a na tworzeniu raportów skończywszy.

System *STATISTICA Data Miner* jest wyposażony w bardzo bogaty wybór metod analitycznych, zarówno wbudowanych w gotowe projekty data mining (szeroki wybór predefiniowanych, gotowych do użycia projektów dla różnego typu zagadnień), jak i dostępnych do budowy własnych projektów. *STATISTICA Data Miner* zawiera pełną funkcjonalność programów: *STATISTICA Pakiet Podstawowy*, *Modele Zaawansowane*, *Analizy Wielowymiarowe*, *Automatyczne Sieci Neuronowe* (z generatorem kodu C), *Trees* i *Analiza Mocy Testów*, składnik IDP, a przede wszystkim kilkanaście specjalnie zaprojektowanych modułów data mining.

Specjalne środowisko graficzne tworzenia projektów analitycznych (patrz poniżej) bazuje na konwencji *przeciągnij i upuść*, dzięki czemu tworzenie projektów data mining jest intuicyjne, a praca z systemem nie wymaga długich przygotowań. Sprawia to, że z systemu *STATISTICA Data Miner* korzystać mogą nawet nowicjusze. Jest on jednocześnie wysoce elastyczny i dostosowywalny. Jak widać poniżej, obszar roboczy w *STATISTICA Data Miner* podzielony jest na cztery sekcje. Podział ten odpowiada czterem etapom projektów data mining.

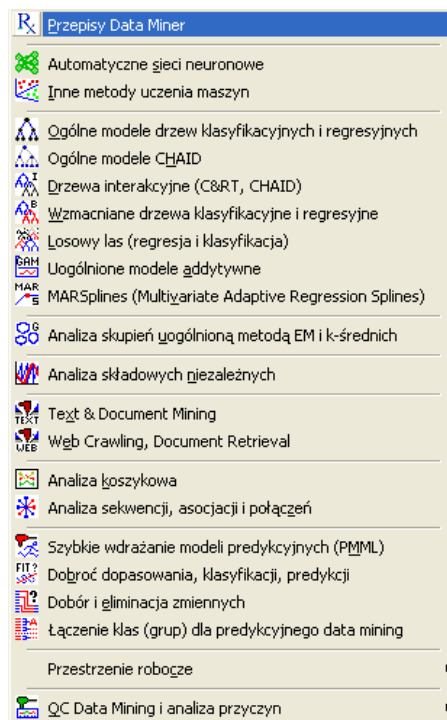


W jednym projekcie data mining mamy możliwość stosowania różnych technik i modeli. Po opracowaniu gotowych modeli *STATISTICA Data Miner* oferuje elastyczne narzędzia do wdrażania i stosowania modeli, zintegrowane ze środowiskiem tworzenia projektów, które umożliwia sterowanie przetwarzaniem danych w obiektach analizy (węzłach) za pomocą poręcznych, zgodnych ze standardami przemysłu informatycznego skryptów *STATISTICA Visual Basic*.

Dane przetwarzane w *STATISTICA* mogą pochodzić z wielu różnych źródeł. System *STATISTICA Data Miner* cechuje łatwość w korzystaniu z danych zewnętrznych. Mamy możliwość podłączenia różnych źródeł danych, tak że jakakolwiek zmiana danych na wejściu powoduje przeliczenie całego projektu, tak aby wyniki odpowiadały aktualnym danym.

Wielkość danych przetwarzanych przez systemy *STATISTICA* jest praktycznie nieograniczona. Specjalne techniki data mining w *STATISTICA Data Miner* są zoptymalizowane pod kątem przetwarzania olbrzymich zbiorów danych.

W jednym projekcie data mining możemy używać różnych technik i modeli. Stosowanie modeli dla nowych danych jest w *STATISTICA Data Miner* bardzo wydajne, dzięki zastosowaniu przenaszalnego, bazującego na XML, języka PMML (*Predictive Models Markup Language*), który można stosować zarówno dla modeli klasyfikacyjnych, jak i regresyjnych. Raz oszacowane modele mogą być później wielokrotnie stosowane dla nowych danych. Możemy także na bieżąco modyfikować ich parametry w miarę napływu danych. Oprócz języka PMML projekty i szacowane modele możemy zapisywać w postaci kodu w językach programowania SVB (zgodnym ze standardowym Visual Basic) oraz C/C++. Dzięki temu analizy przeprowadzane w *STATISTICA Data Miner* mogą być wykorzystywane także w innych aplikacjach. Najczęstszym jednak zastosowaniem dla modeli zapisywanych w tych różnych językach jest ich ponowne wykorzystanie w projektach data mining w *STATISTICA Data Miner*. Zapisane wcześniej analizy mogą być wstawiane do przestrzeni projektowej jako nowe węzły i łączone z pozostałymi elementami projektu.



Nowością w *STATISTICA Data Miner* są *Przepisy Data Miner* – kompletne rozwiązanie, dzięki któremu tworzenie modeli predykcyjnych staje się systematycznym procesem wykonywanym krok po kroku według ustalonej procedury. Celem *Przepisów Data Miner* jest umożliwienie łatwego i szybkiego wydobywania z danych użytecznych informacji. W *Przepisach* użytkownik tworzy model, wykonując krok po kroku pewien ogólny przepis obejmujący przygotowanie i oczyszczenie danych, a następnie użycie ich do stworzenia modeli predykcyjnych i zastosowanie modeli dla nowych obiektów. Wiele domyślnych ustawień zależy od danych, np. jeśli w pliku mamy wiele zmiennych, to musimy zastosować dobór zmiennych do modelowania, a jeśli zmiennych jest niewiele, to dobór zmiennych jest opcjonalny. Domyślne ustawienia systemu zazwyczaj dają zadowalające wyniki, ale zaawansowani użytkownicy mają do dyspozycji dodatkowe ustawienia i mogą zmieniać domyślne ustawienia.

*STATISTICA Data Miner* zawiera kilkanaście specjalnie zaprojektowanych modułów data mining:

- ◆ **Interakcyjne drażenie danych:** Zadaniem tego modułu jest umożliwienie jednoczesnego tworzenia tabel i wykonywania graficznej, eksploracyjnej analizy danych, tak aby można było szybko przyrzeć się rozkładowi badanych zmiennych, ich związkom z innymi zmiennymi i wykrycia obserwacji należących do konkretnych grup.
- ◆ **Automatyczne sieci neuronowe** to kompletne, korzystające z technik sztucznej inteligencji narzędzia modelowania i prognozowania. Metody dostępne w *STATISTICA Automatyczne Sieci Neuronowe* znajdują tu zastosowanie w projektach data mining.
- ◆ **Analiza składowych niezależnych** służy do filtrowania sygnałów za pomocą dobrze sprawdzonej i niezawodnej techniki statystycznej znanej jako *Independent Component Analysis (ICA)*. Zastosowana implementacja tej metody pozwala rozwiązywać



problemy separacji sygnałów metodami jednoczesnego wyodrębniania i deflacji, wstępnie przetwarzać dane, a także stosować rozwiązanie dla nowych danych.

- ◆ **Analiza skupień uogólnioną metodą EM i k-średnich** to metody uczenia bez nauczyciela (nieukierunkowane), stosowane do rozpoznawania wzorców i segmentacji, dla których można jednak zastosować metody wdrażania zazwyczaj wykorzystywane w modelach predykcyjnych. Skupienia mogą być definiowane w oparciu o wartości zmiennych ciągłych (liczbowych, ilościowych) i skategoryzowanych (jakościowych).
- ◆ **Analiza koszykowa** jest implementacją tzw. „algorytmu wykrywania a priori” (poszukiwania) reguł powiązania typu: „klienci, którzy zamówili produkt A, często zamawiają produkt B lub C”, „pracownicy, którzy dobrze oceniają rozwiązanie X, również często narzekają na rozwiązanie Y, ale zadowolą ich rozwiązanie Z”. Moduł ten umożliwia przetwarzanie olbrzymich zbiorów danych w poszukiwaniu powiązań (współzależności), z wykorzystaniem zdefiniowanych wstępnie wartości progowych do wykrywania.
- ◆ **Ogólne modele drzew klasyfikacyjnych i regresyjnych (G-Trees)**: Moduł ten zawiera kompletną implementację metod CART®. Program ma bardzo wiele opcji sterujących budową drzew, przycinania, jak i wyboru najlepiej dopasowanego rozwiązania.
- ◆ **Ogólne modele CHAID** (*Chi-squared Automatic Interaction Detector*) to jedna z najstarszych metod drzew klasyfikacyjnych. CHAID buduje drzewa, w których z węzłów mogą wychodzić więcej niż dwie gałęzie i wykonuje to za pomocą algorytmu, który nadaje się zwłaszcza do analizy dużych zbiorów danych.
- ◆ **Drzewa interakcyjne (C&RT, CHAID)** W tym module budowane są (rosną) automatycznie lub za pomocą algorytmu sterowanego regułami i kryteriami określonymi przez użytkownika, poprzez interaktywny, graficzny interfejs, lub obydwoma metodami.
- ◆ **Drzewa wzmacniane**: Moduł ten pozwala zastosować wzmacnianie (*boosting*) do zadań predykcyjnego data mining. Użytkownik może sterować procedurą budowy modelu, na każdym jej etapie dostępne są szczegółowe podsumowania, pozwalające obserwować i oceniać postęp uzyskiwany na każdym kroku estymacji.
- ◆ **Losowy las** (regresja i klasyfikacja) służy do budowy złożonych modeli klasyfikacji i regresji. Jest szczególnie odpowiedni dla skrajnie dużych liczb predyktorów. *STATISTICA Losowy las* jest realizacją algorytmu Random Forest opracowanego przez Breimana. *STATISTICA Losowy las* daje użytkownikowi pełną kontrolę nad kluczowymi ustawieniami procedury dopasowania modelu.
- ◆ **Uogólnione modele addytywne** to implementacja metod rozwiniętych i spopularyzowanych przez Hastiego i Tibshiraniego. W programie można stosować wiele rozkładów dla zmiennej zależnej. Podobnie wiele funkcji wiążących (ang. *link function*) może być użytych do opisu zależności zmiennej zależnej od predyktorów.
- ◆ **MARSplines** (*Multivariate Adaptive Regression Splines* – Wielozmienna regresja adaptacyjna z użyciem funkcji sklepanych) jest implementacją uogólnienia techniki służącej do rozwiązywania zarówno problemów regresyjnych, jak i klasyfikacyjnych,



w których celem jest znalezienie wartości zmiennych wyjściowych (zależnych) na podstawie zmiennych wejściowych (predykcyjnych).

- ◆ **Inne metody uczenia maszyn** (*Machine Learning*) to zestaw zaawansowanych metod modelowania dla zagadnień regresyjnych i klasyfikacyjnych przy wielu zmiennych niezależnych (predyktorach) i zależnych. Te metody to:
  - wektory nośne (*Support Vector Machines*),
  - naiwny klasyfikator Bayesa,
  - K-najbliższych sąsiadów (*K-Nearest Neighbors*).
- ◆ **Szybkie wdrażanie modeli predykcyjnych**: W tym module można szybko wygenerować predykcje jednego lub wielu uprzednio zbudowanych modeli zapisanych w standardowym formacie PMML (*Predictive Model Markup Language*).
- ◆ **Dobroć dopasowania, klasyfikacji, predykcji** (*Goodness of fit*). Przeznaczeniem tego modułu jest ocena modeli predykcyjnych, zarówno zmiennych zależnych ciągłych, jak i skategoryzowanych. Obliczane są różne statystyki dobroci dopasowania na podstawie przewidywanych i obserwowanych wartości lub klasyfikacji. Powstaje też wiele podsumowujących wykresów ilustrujących wyniki badania.
- ◆ **Dobór i eliminacja zmiennych** umożliwia przeprowadzanie automatycznego doboru podzbiorów zmiennych ze skrajnie dużych zbiorów danych lub baz danych, połączonych w celu zdalnego przetwarzania (IDP). Może on operować na praktycznie nieograniczonej liczbie zmiennych; można sprawdzić dosłownie miliony zmiennych wejściowych w celu wyszukania predyktorów dla zagadnień regresyjnych lub klasyfikacyjnych.
- ◆ **Łączenie klas (grup) do predykcyjnego data mining** jest metodą przetwarzania predyktorów jakościowych o dużej liczbie kategorii.

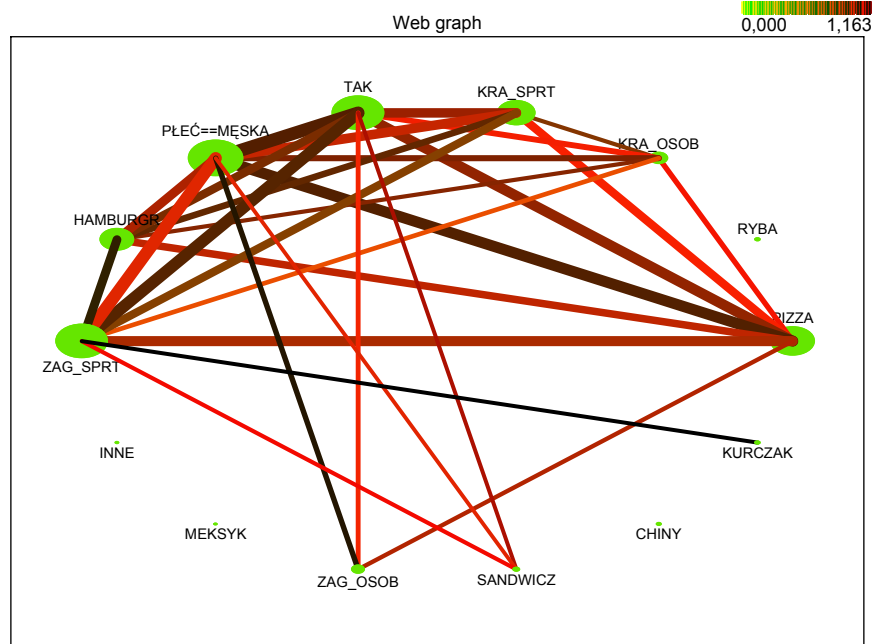


### **STATISTICA Sequence, Association and Link Analysis (SAL) PL**

*STATISTICA SAL* wykorzystuje szereg najnowszych technik analitycznych przeznaczonych do znajdowania w zbiorach danych reguł, jakim podlegają „koszyki zakupów”. Program ten można traktować jako rozszerzenie technik dostępnych w *Analizie koszykowej* zawartej w *STATISTICA Data Miner*. *STATISTICA SAL* pomaga w następujących analizach:

- ◆ wykrywanie reguł asocjacji (ustalenie reguł, które produkty kupowane są razem),
- ◆ analiza sekwencji (badanie, w jakiej kolejności zjawiska zachodzą po sobie),
- ◆ analiza skojarzeń (dzięki niej można, np. analizując dane dotyczące klientów, wnioskować o tym, co będą oni chcieli kupić, bazując na informacjach o wcześniejszych ich zakupach).

*STATISTICA SAL* stanowi doskonałe rozszerzenie pakietu *STATISTICA Data Miner*, może też działać samodzielnie.



### STATISTICA Text Miner

STATISTICA Text Miner jest rozszerzeniem systemu STATISTICA Data Miner, służącym do przekształcania pozbawionych struktury tekstów w użyteczną wiedzę, wspomagającą podejmowanie decyzji.

Document Retrieval: Arkusz2

Level of depth: 14 Max. items in crawling tree: 500 Domain restricted

Destination (URL or Folder): Folder Browser Add To Crawl Options

File Filter: Web Pages (\*.html;\*.htm)

Crawling tree

```

s://serwer/www.statsoft.pl/products/./skrypty
s://serwer/www.statsoft.pl/products/desktop.h
s://serwer/www.statsoft.pl/products/dataminer
s://serwer/www.statsoft.pl/products/webserve
s://serwer/www.statsoft.pl/products/enterprise
erwer/www.statsoft.pl/dms.html
s://www.statsoft.com/products/documentmana
erwer/www.statsoft.pl/products/mspc.htm
s://serwer/www.statsoft.pl/products/mspcdeta
http://www.statsoft.com/products/mspcdeta
    
```

Start Stop Delete

Start & put the result directly to a spreads...  
Start & put the result directly to local fol...

Folder Option  
 Save only filtered file results when cr...

Complete <http://www.statsoft.com/products/po>

TM results: Arkusz3 w Skoroszyt1

Number of documents: 14  
Number of selected words: 210  
Number of unselected words: 0

Statistic for occurrence  
 Frequency  Inverse document frequency  
 Binary frequency  Log frequency

Word	Count	Files	Stemmed	Status
sieci	220	11	sieci	Selected
sigma	52	11	sigma	Selected
six	51	11	six	Selected
skali	11	11	skali	Selected
specjalnie	35	12	specjalnie	Selected
statistica	1355	13	statistica	Selected
statsoft	105	12	statsoft	Selected
stosowanie	40	11	stosowanie	Selected

Quick Words SVD Search Save results

Perform SVD (singular value decomposition) Current results for singular value are computed for: Frequency.

Scree plot  Singular values  
 Word coefficients  Document scores

Number of singular values considered: 10

Sum of squares of word residuals  
 Word importance from SVD analysis  
 Save SVD results for deployment



W projektach data mining wykorzystuje się dane o różnej postaci – nie zawsze mają one strukturę umożliwiającą łatwe wykonanie analizy. *STATISTICA Text Miner* wglębia się w dane, które nie mają tradycyjnej struktury „arkusza danych”, i odnajduje w nich użyteczną informację. *STATISTICA Text Miner* umożliwia m.in. badanie zawartości witryn internetowych, dużych repozytoriów dokumentów, budowanie modeli predykcyjnych na podstawie nieobrobionych tekstów i danych innego typu.

## Systemy korporacyjne

Każde z wyżej opisanych narzędzi analitycznych *STATISTICA* może działać na pojedynczym komputerze lub w systemach sieciowych, na wielu stacjach roboczych. W takim przypadku każdy z analityków ma do dyspozycji odpowiedni zestaw narzędzi analitycznych. Łącząc wszystkie lub wybrane narzędzia analityczne z jednym z niżej opisanych systemów korporacyjnych, otrzymujemy rozwiązanie, które nie tylko pozwala na korzystanie z odpowiednich narzędzi analitycznych przez wielu użytkowników, ale także organizuje wspólną pracę różnych osób. W szczególności systemy korporacyjne *STATISTICA* pozwalają wielu użytkownikom korzystać ze wspólnych szablonów zapytań do baz danych, analiz, raportów przechowywanych w jednym miejscu z zapewnieniem kontroli dostępu itp.



### *STATISTICA Enterprise PL*

*STATISTICA Enterprise* to system analizy danych ogólnego przeznaczenia, m.in. dla zastosowań typu *business intelligence* w marketingu, finansach oraz w innych dziedzinach. Użytkownicy *STATISTICA Enterprise* mogą korzystać ze wspólnych, nawet bardzo złożonych zapytań, pozwalających na przeglądanie wybranych danych z repozytoriów centralnych, oraz współużytkowanie centralnie uaktualnianych skryptów analiz (np. predefiniowane raporty dostosowywane przez administratorów lub analityków). *STATISTICA Enterprise* ma własny (niezależny od systemu operacyjnego) system użytkowników i grup z możliwością zarządzania uprawnieniami do zasobów systemu, narzędzie do definiowania szablonów raportów oraz generator raportów (raporty generowane automatycznie co pewien czas lub o określonym czasie).

*STATISTICA Enterprise* zorganizowany jest wokół centralnej hurtowni danych, którą można zainstalować na wszystkich popularnych systemach zarządzania bazami danych, w tym na popularnych skalowalnych systemach takich jak Sybase, Oracle, MS SQL Server czy IBM DB2. W skład systemu *STATISTICA Enterprise* wchodzi narzędzia administrowania systemem (w pełni zintegrowane z innymi funkcjami) oraz hurtownia danych, w której gromadzone są dane *STATISTICA*. System uprawnień pozwala administratorowi przypisywać użytkownikom prawa dostępu do konkretnych danych.

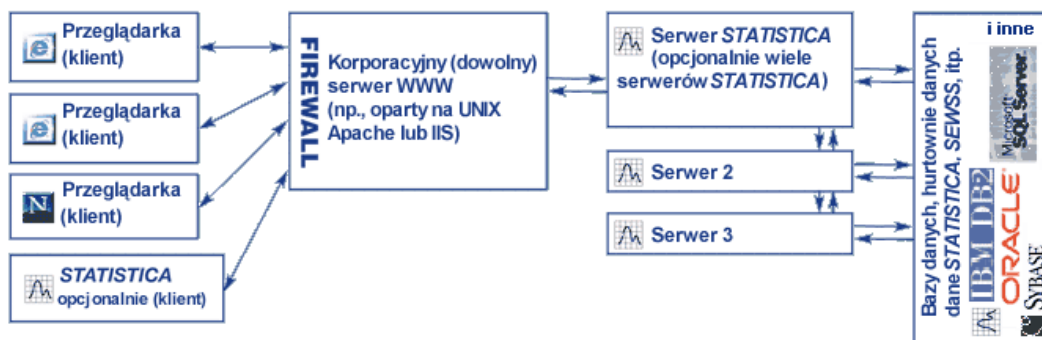
Wyniki pracy mogą być udostępniane w środowisku lokalnym (innym użytkownikom *STATISTICA Enterprise* posiadającym odpowiednie uprawnienia) oraz rozpowszechniane w szerszym zakresie (np. jako raporty publikowane w Internecie lub Intranecie).



## WebSTATISTICA Server

jest w pełni skalowalnym internetowym systemem udostępniania i analizy danych w skali całej organizacji. Umożliwia on zdalne uruchamianie *STATISTICA* w ramach przeglądarki internetowej (wszystkie obliczenia są wykonywane na serwerze), a także szybki i łatwy dostęp do danych, raportów i narzędzi analitycznych z praktycznie każdego komputera połączanego z Internetem. Pozwala to na współpracę osób z tego samego biura, jak i z różnych kontynentów.

System ten w pełni wykorzystuje wielowarstwową architekturę klient-serwer. *WebSTATISTICA* wykorzystuje przetwarzanie rozproszone i wielowątkowość, co pozwala optymalnie wykonywać bardzo duże zadania obliczeniowe. Technologia ta umożliwia szybkie wykonanie nawet bardzo dużych projektów, z pełnym wykorzystaniem wielu procesorów serwera lub wielu serwerów pracujących równolegle.



*WebSTATISTICA* może współpracować z wieloma popularnymi serwerami WWW (np. Apache lub IIS) i istniejącymi protokołami zabezpieczeń (*firewall*), zgodnie z wymogami korporacyjnymi. Z systemu można korzystać z każdego komputera na świecie połączanego z Internetem, wyposażonego w przeglądarkę internetową. Nie ma także ograniczeń związanych z systemem bazy danych i platformą, na której on pracuje. *WebSTATISTICA Server* może pracować na danych pochodzących ze wszystkich popularnych systemów baz danych.

System *WebSTATISTICA* jest oferowany jako kompletne rozwiązanie, zawierające funkcjonalność poszczególnych produktów z rodziny *STATISTICA* lub ich kombinacji, począwszy od *STATISTICA Pakiet Podstawowy*, a skończywszy na systemach *STATISTICA Enterprise* czy *PROCEED*. Rozwiązanie to jest gotową do instalacji aplikacją, z w pełni funkcjonalnym graficznym środowiskiem użytkownika. Jednocześnie zawiera ono narzędzia programistyczne umożliwiające dostosowywanie wszystkich elementów systemu i jego rozbudowę.



## Zarządzanie dokumentami



### *STATISTICA Document Management System*

to kompletny, skalowalny system do zarządzania wszelkiego typu dokumentami zapisanymi w formie elektronicznej (nie tylko dokumentami *STATISTICA*, ale również innymi dokumentami używanymi w organizacji). Produkt ten umożliwia sprawne, wydajne i bezpieczne operowanie dokumentami różnych typów, szybkie odnajdywanie dokumentów, udostępnianie, przeszukiwanie ich zawartości, a także przeglądanie dokumentu, jego edycję (rejestrując przy tym proces edycji i odpowiednio zarządzając kolejnymi wersjami) itp.

*STATISTICA Document Management System* pozwala na zapisanie dokumentu do bezpiecznej bazy danych. Po każdej edycji dokumentu tworzona jest jego nowa wersja, z którą powiązane są metadane zawierające informację o czasie i osobie wprowadzającej zmiany. W trakcie edytowania dokumentu tworzona jest jego nowa wersja, której nadawany jest odpowiedni certyfikat autentyczności i podpis elektroniczny. Dokumenty i ich nowe wersje mogą wymagać zatwierdzenia przez odpowiednie osoby, z podaniem okoliczności i celu edycji dokumentu. Dokumenty można także zabezpieczyć przed jakąkolwiek dalszą edycją, bądź tworzeniem nowych wersji.

Produkt ten jest zaprojektowany z myślą o zapewnieniu zgodności z uregulowaniami norm ustalających sposób zarządzania dokumentami, np. FDA 21 CFR Część 11, czy ISO 9000, 9001, 1400.

## StatSoft Polska

StatSoft Polska jest działającym od 1995 roku biurem amerykańskiej firmy StatSoft Inc., która ma już ponad dwudziestoletnie doświadczenie w zakresie systemów analitycznych i związanych z nimi usług. Główną domeną działalności StatSoft Polska jest szeroko rozumiana analiza danych. Profesjonalne oprogramowanie analityczne *STATISTICA* oraz kompleksowe usługi doradcze i szkoleniowe realizowane przez doświadczonych specjalistów stanowią gwarancję zadowolenia klientów. Niniejszy artykuł jest poświęcony rozwiązaniom z rodziny *STATISTICA*.

StatSoft uzupełnia ofertę w zakresie analizy danych o szkolenia i usługi konsultingowe. Kursy organizowane przez StatSoft Polska przeznaczone są dla wszystkich, którzy w krótkim czasie chcą zdobyć lub pogłębić wiedzę w zakresie przetwarzania i analizy danych (statystycznej, data mining) w różnych dziedzinach i zastosowaniach oraz obsługi programów z rodziny *STATISTICA*. StatSoft Polska organizuje regularnie odbywające się szkolenia otwarte w Krakowie (obecnie jest już ponad 40 różnych tematów do wyboru), a także kursy na zamówienie, które są dostosowywane do indywidualnych potrzeb zamawiającego. Nasze kursy są kursami autorskimi, a nie kalkami szkoleń przygotowanych za granicą i przetłumaczonych na język polski. Dzięki temu, że programy naszych kursów są przygotowywane od początku do końca przez prowadzących, specjalistów z zakresie



analizy danych w różnych zastosowaniach, są one doskonale dopasowane do polskich realiów i spełniają oczekiwania uczestników. Potwierdzeniem najwyższej jakości szkoleń StatSoft Polska są opinie uczestników, które zamieszczone są na stronie internetowej [www.StatSoft.pl/kursy.html](http://www.StatSoft.pl/kursy.html). Są tam też szczegółowe informacje o kursach.

StatSoft Polska oferuje usługi konsultingowe w zakresie analizy danych i budowy systemów analitycznych. Mogą one obejmować bardzo szeroki zakres zagadnień, np. analizę zgromadzonych danych, budowę modeli analitycznych, opracowanie i implementację kompletnych systemów analitycznych (od zbierania i gromadzenia danych, przez zaprojektowanie szablonów analiz, po automatyczne generowanie raportów), doradztwo w zakresie gromadzenia i analizy danych, tworzenie dedykowanych aplikacji działających w środowisku *STATISTICA* i wykorzystujących jej unikalne właściwości analityczne. Do StatSoft Polska może zwrócić się każdy, kto chce zasięgnąć porady w zakresie analizy danych lub chce skorzystać z pomocy naszych doświadczonych specjalistów. Więcej informacji o usługach konsultingowych StatSoft Polska, np. opinie o wybranych wdrożeniach, można znaleźć na stronie [www.statsoft.pl/konsulting.html](http://www.statsoft.pl/konsulting.html).