



## **PROGRAMY Z RODZINY *STATISTICA* JAKO NARZĘDZIA ANALIZY DANYCH**

*dr Janusz Wątroba<sup>1</sup>*

Programy wchodzące w skład rodziny *STATISTICA* to uniwersalne, zintegrowane systemy przeznaczone do zarządzania danymi, statystycznej analizy danych oraz tworzenia wykresów. Obejmują one niezwykle obszerny zestaw prostych i zaawansowanych procedur analitycznych znajdujących zastosowanie do analizy zjawisk i procesów badanych w nauce, biznesie i technice. Zawierają nie tylko procedury statystyczne i graficzne ogólnego zastosowania, ale także inne specjalistyczne techniki analityczne. Rodzina programów *STATISTICA* jest stale rozwijana i to zarówno w sensie rozbudowy możliwości analitycznych jak i pod względem stosowania coraz to nowszych technologii informatycznych. Dzięki temu znajduje zastosowanie wszędzie tam, gdzie z danych, poprzez odpowiednią analizę należy wydobyć istotne informacje i przedstawić je w postaci tabel z wynikami i wykresów ułatwiających interpretację prawidłowości dotyczących badanego fragmentu rzeczywistości bądź wspomoczenie procesu decyzyjnego.

### **StatSoft Inc. i StatSoft Polska**

Firma StatSoft Inc., producent oprogramowania z rodziny *STATISTICA* powstała w 1984 roku. Jej siedziba mieści się w miejscowości Tulsa w USA. Obecnie StatSoft Inc. jest jednym z największych producentów oprogramowania do analiz statystycznych i wizualizacji na świecie. StatSoft Polska Sp. z o.o. jest polskim biurem StatSoft Inc. i zajmuje się dystrybucją i sprzedażą produktów firmy StatSoft w Polsce, pomocą techniczną, prowadzeniem kursów obsługi programów z rodziny *STATISTICA* i kursów statystycznej analizy danych oraz wdrażaniem oferowanych systemów analizy danych. Oferujemy także usługi konsultingowe z zakresu badań, przetwarzania i statystycznej analizy danych z różnych dziedzin (np. w biomedycynie, badaniach opinii publicznej, badaniach marketingowych, zagadnieniach statystycznego sterowania jakością).

### **Kompleksowa oferta systemów do analizy danych z rodziny *STATISTICA***

Biorąc pod uwagę przeznaczenie analiz, sposoby dostępu do danych oraz ilość użytkowników, którzy mają mieć dostęp do narzędzi analitycznych systemy analizy danych firmy StatSoft można podzielić na trzy grupy. Pierwszą grupę stanowią programy przeznaczone do szeroko rozumianej analizy danych dla potrzeb pojedynczego użytkownika lub grupy użytkowników. Druga grupa to programy przeznaczone do analiz związanych z zagadnieniem zapewnienia jakości. I wreszcie trzecia grupa to zintegrowane systemy analizy danych przeznaczone dla organizacji w których

---

<sup>1</sup> StatSoft Polska Sp. z o.o.



istnieje potrzeba przeprowadzania analiz na dużych zbiorach czy bazach danych lub w przypadku potrzeby stworzenia kompleksowego systemu sterowania jakością w skali całego przedsiębiorstwa.

### **STATISTICA - narzędzia analityczne**

**STATISTICA.** W tej grupie produktów podstawowym programem jest pakiet *STATISTICA*. Od kilku lat niezmiennie otrzymuje najwyższe oceny w porównaniach i testach przeprowadzanych na całym świecie przez niezależnych ekspertów. Swoją sukces *STATISTICA* zawdzięcza przede wszystkim łatwości obsługi, bogactwu procedur statystycznych, doskonałej grafice i zastosowaniu najnowszych technologii informatycznych oraz procedur numerycznych. Program zawiera obszerny zestaw procedur analitycznych wykorzystywanych w badaniach naukowych (w zagadnieniach biomedycznych, społecznych, inżynierskich, ekonomicznych, marketingowych) oraz w biznesie i przemyśle. Umożliwia kompletną obróbkę danych, począwszy od wprowadzania i weryfikacji danych do momentu utworzenia raportu. Program w pełni wykorzystuje możliwości środowiska Windows (95, 98, 2000 i NT). Szczególną uwagę zwrócono na ułatwienie sposobu komunikacji z użytkownikiem. Może on być dostosowywany do jego potrzeb na wiele sposobów. Program ma bardzo rozbudowane możliwości przyspieszania i automatyzacji zadań. Staranny dobór algorytmów obliczeniowych oraz zoptymalizowane mechanizmy zarządzania pamięcią zapewniają niespotykaną w innych tego typu programach szybkość i dokładność obliczeń przy niezbyt wygórowanych wymaganiach sprzętowych. Program oferuje prawie nieograniczone możliwości graficznej prezentacji danych surowych i wyników obliczeń oraz łatwy dostęp do modyfikacji wszystkich elementów utworzonego wykresu. Twórcy programu zadbali o pełną integrację metod analitycznych i graficznych. Program posiada obszerną dokumentację, zarówno w postaci drukowanej jak i w formie elektronicznego podręcznika wykorzystującego architekturę hipertekstową. Program *STATISTICA* posiada budowę modułową. Wszystkie moduły stanowią elementy zintegrowanego pakietu.

**Quick STATISTICA.** Kolejny produkt w tej grupie to *Quick STATISTICA*. Jest to program oferujący podstawowe metody statystycznej analizy danych (np. obliczanie statystyk opisowych, testy nieparametryczne, tworzenie i analiza tabel wielodzzielczych, analiza wariancji i analiza regresji) oraz pełne możliwości graficzne i zarządzanie danymi. Jest to zatem podzbiór pakietu *STATISTICA* nie zawierający bardziej zaawansowanych procedur statystycznych.

**STATISTICA Neural Networks.** Jest to program poszerzający możliwości analityczne programów z rodziny *STATISTICA* o rozwiązania wykorzystujące technikę sieci neuronowych, które mogą stanowić potężne narzędzia zgłębiania danych (data mining) oraz mogą być stosowane w różnych dziedzinach, takich jak finanse, badania inżynierskie czy biomedycyna. Program może być wykorzystywany jako samodzielna aplikacja lub w połączeniu z programem *STATISTICA*.

**STATISTICA Power Analysis.** Jest to program zawierający narzędzia do analizy mocy testów, wyznaczania liczebności próby oraz estymacji przedziałowej. Zagadnienia mocy testów i estymacji odpowiedniej liczebności próby mają bardzo duże praktyczne znaczenie, gdyż już na etapie projektowania badań pomagają lepiej wykorzystać dostępne środki. Jako jedyny spośród dostępnych na rynku programów tego typu *STATISTICA Power Analysis* wychodzi poza testy „braku różnic” i pozwala zastosować nowoczesne techniki opierające się na estymacji przedziałowej.

### **STATISTICA dla jakości**

**STATISTICA Quality Control Charts.** Jest to bardzo obszerny zestaw narzędzi przeznaczonych do sterowania jakością. Program bardzo dobrze nadaje się do zastosowania na hali produkcyjnej do monitorowania w czasie rzeczywistym procesów o dowolnym typie i złożoności. Narzędzia w nim



zawarte są szczególnie użyteczne w przypadku ich integracji z istniejącymi systemami akwizycji danych.

**STATISTICA Design of Experiments.** Program ten oferuje szeroki zestaw metod przeznaczonych do planowania i analizy doświadczeń. Może być z powodzeniem stosowany w przemyśle (np. kreowanie jakości), naukach rolniczych i biomedycznych. Do pracy wymaga jednego z trzech programów: *STATISTICA*, *Quick STATISTICA* lub *STATISTICA Quality Control Charts*.

**STATISTICA Process Analysis.** Jest to program przeznaczony głównie do zastosowań w przemyśle. Zawiera narzędzia wykorzystywane w analizie wydolności procesu oraz analizie powtarzalności i odtwarzalności pomiarów.

### Zintegrowane systemy analizy danych

**STATISTICA Enterprise Wide System (SENS).** Jest to kompleksowy system analizy danych o ogólnym przeznaczeniu. Tworzą go narzędzia analityczne i graficzne *STATISTICA*, wbudowane mechanizmy dostępu do zewnętrznych baz danych oraz narzędzia pracy grupowej zintegrowane wokół dedykowanej hurtowni danych. System udostępnia również narzędzia do publikacji i rozpowszechniania informacji i wyników analiz. Współpracuje ze wszystkimi popularnymi na rynku bazami danych (np. Oracle, Sybase, IBM DB2, SQL Server, Informix i inne). Może współpracować z istniejącymi systemami informatycznymi (np. ERP).

**STATISTICA Enterprise Wide SPC System (SEWSS).** Jest to kolejny system analizy danych umożliwiający kompleksowe stosowanie statystycznego sterowania jakością w skali całego przedsiębiorstwa. Może być wykorzystywany zarówno w małych laboratoriach, na pojedynczej linii produkcyjnej jak również w dużych przedsiębiorstwach i firmach. Posiada bardzo rozbudowane mechanizmy obsługi środowiska wprowadzania danych i monitorowania procesu. Pozwala zbudować w skali całej organizacji całościowe rozwiązania zagadnienia analizy danych i statystycznego sterowania jakością.

### Wybrane zalety oprogramowania firmy StatSoft

Programy z rodziny *STATISTICA* przebojem zdobyły rynek oprogramowania do analiz statystycznych i wizualizacji danych. Oto zestawienie najważniejszych zalet naszych produktów:

- ◆ możliwość przeprowadzania kompletnej analizy danych; od wprowadzenia danych w specjalnym arkuszu lub ich pobrania z innych aplikacji (np. baz danych w dowolnym formacie) poprzez wszechstronną analizę danych i bogatą wizualizację wyników oraz analizę typu „Co - jeśli”, do tworzenia raportów z analiz i możliwości łatwego ich rozpowszechniania np. w Intranecie lub Internecie
- ◆ na każdym etapie analizy istnieje możliwość przedstawiania pośrednich i końcowych wyników w postaci czytelnych tabel i wykresów ułatwiających merytoryczną interpretację uzyskanych rezultatów
- ◆ możliwość łatwej i szybkiej modyfikacji analizy i porównywania wyników uzyskanych w oparciu o różne zbiory danych (w przypadku nowych modułów można dany moduł otwierać wielokrotnie)
- ◆ dostosowanie interfejsu użytkownika zarówno do potrzeb początkujących użytkowników jak i użytkowników bardziej zaawansowanych (dostęp do wbudowanych języków *SCL* i *STATISTICA Basic* oraz składni *VGLM*, *VGLZ*, *VGSR* oraz *VPLS*)
- ◆ łatwa wymiana informacji z innymi aplikacjami środowiska Windows



- ◆ bardzo dobra relacja oferowanych możliwości analitycznych do ceny już w pakiecie podstawowym
- ◆ możliwość rozbudowy możliwości programów poprzez dokupienie modułów przeznaczonych do zastosowań specjalnych (np. *STATISTICA Neural Networks* lub *STATISTICA Process Analysis*)
- ◆ możliwość integracji systemów analizy danych z rodziny *STATISTICA* z istniejącymi w danej organizacji systemami informatycznymi
- ◆ stały, nieprzerwany rozwój programów z rodziny *STATISTICA* polegający na poszerzaniu możliwości analitycznych i graficznych oraz wprowadzaniu nowych technologii informatycznych
- ◆ oprócz oprogramowania firma StatSoft proponuje bogatą ofertę szkoleń i kursów obsługi programów z rodziny *STATISTICA* oraz kursów statystycznej analizy danych w różnych dziedzinach zastosowań.

Własności programów z rodziny *STATISTICA* powodują, że w efekcie użytkownik otrzymuje do dyspozycji doskonale merytorycznie i bardzo przyjazne środowisko do przeprowadzania różnego rodzaju analiz pomocnych w opracowywaniu danych pochodzących z badań naukowych. Bogactwo możliwych do wykorzystania metod analizy powoduje, że programy te doskonale sprawdzają się we wszystkich tych dyscyplinach naukowych w których metody analizy statystycznej są powszechnie stosowane (np. w naukach społecznych, ekonomicznych czy biomedycznych). Najlepszą rekomendacją dla naszych programów może być wynik obszernego przeglądu porównawczego zamieszczonego w czasopiśmie *NATURE* (z lipca 1998 roku) oraz niezwykle pochlebne opinie wyrażone przez przedstawicieli środowiska naukowego w innym renomowanym czasopiśmie naukowym - *SCIENCE* (w numerze z maja 1998 i 1999).