



StatSoft®

business intelligence • data mining • sterowanie jakością • analizy w środowisku internetowym

STATISTICA Enterprise-wide SPC System (SEWSS):
Narzędzia analizy danych
dla systemów zbierania i przechowywania danych

Część 1: Współpraca z bazami danych

Praktyczny materiał informacyjny StatSoft

Wstęp

Firma StatSoft, producent zestawu programów **STATISTICA**, współpracuje ze swoimi klientami projektując i dostarczając im rozwiązania najwyższej jakości i niezawodności. **STATISTICA Enterprise-wide SPC System (SEWSS)** jest najbardziej zaawansowanym systemem analizy danych dla działów badań i rozwoju (R&D) jak i zapewnienia jakości. Z doświadczeń StatSoft, zebranych przy współpracy z klientami wynika, że głównymi warunkami stosowana analizy danych są:

- **Łatwe korzystanie z systemu.** Bez względu na to jak wielkie są możliwości narzędzia analizy danych, nie będzie ono używane przez użytkowników, jeśli nie jest intuicyjne w obsłudze.
- **Korzystanie z różnych źródeł danych.** Pracownicy, którzy chcą analizować dane i wiedzą jakie informacje są im potrzebne, często muszą poświęcić dużo czasu na współpracę z osobami, które dokładnie znają struktury baz danych i sposoby pobierania danych. System powinien zminimalizować czas i wysiłek potrzebny na pobranie danych, poprzez pełną integrację i przejrzystą współpracę z repozytoriami danych.
- **Współpraca.** Zadania działów badania i rozwoju (R&D) i zapewnienia jakości rzadko realizuje jedna osoba działająca samodzielnie. Najczęściej projekty takie wymagają wspólnej pracy wielu osób.

Biorąc pod uwagę powyższe fakty, firma StatSoft zaprojektowała system **SEWSS**, który wprowadzono do sprzedaży w 1997 r. Od tego czasu system jest ciągle rozwijany i udoskonalany.

SEWSS stanowi rozbudowany zestaw narzędzi data mining, analiz i wizualizacji danych, przyczym wszystkie one zintegrowane są w jeden system. Użytkownicy **SEWSS** korzystają ze środowisk dostosowanych do przypisanych im zadań (np. inżynier jakości). W systemie można definiować szablony analiz umożliwiające automatyzację rutynowych zadań. W ramach systemu, wielu użytkowników może wspólnie korzystać z danych, raportów i wyników analiz przy zachowaniu bezpieczeństwa informacji. Każdy pracownik, technik i analityk może potrzebować dostępu do danych i narzędzi analitycznych. **SEWSS** jest rozwiązaniem zaspokajającym te potrzeby.

Niniejszy artykuł jest pierwszym z serii trzech tekstów wprowadzających do kluczowych właściwości systemu **SEWSS**, ze szczególnym uwzględnieniem narzędzi współpracy z istniejącymi źródłami danych. Korzyści, jakie dają te funkcje, to przede wszystkim zapewnienie organizacji możliwości pełnego i nieustającego wykorzystania wartościowych danych, zbieranych w obrębie organizacji.

Przeznaczenie **SEWSS**

W wielu organizacjach bieżące wykorzystanie bogactwa zbieranych i przechowywanych danych nie jest optymalne. Powszechne jest przeświadczenie, że potrzebny jest zintegrowany, kompleksowy system analizy danych i raportowania. Przeświadczeniu temu towarzyszy obawa, że opracowanie i wdrożenie takiego systemu byłoby kosztowne, a potencjalni użytkownicy musieliby zostać przeszkoleni, aby z niego skorzystać.

Wiele organizacji dysponuje bogactwem danych, a ich pracownicy mogliby korzystać z tych danych przy podejmowaniu decyzji. Niestety brak odpowiednich narzędzi (spełniających przedstawionej powyżej wymogi) oraz ograniczenia czasowe, uniemożliwiają czerpanie korzyści z analizy danych. Potrzebny jest zintegrowany, kompleksowy system dysponujący narzędziami budowy zapytań i pobierania danych, analizy i wizualizacji, a wszystko to w bezpiecznym, wieloużytkownikowym środowisku pracy zespołowej.

Wyobraźmy sobie system, w którym do uzyskania wyników analizy aktualnych danych wystarczy:

- 1 - być uprawnionym do korzystania z danej analizy
- 2 - kliknąć jej nazwę na hierarchicznie uporządkowanej liście dostępnych analiz.

Użytkownik nie musi wiedzieć, gdzie i w jaki sposób przechowywane są dane. Nie musi on także wiedzieć, że np. dane pochodzą z dwóch różnych źródeł, że specyfikacje produktu zostały automatycznie odczytane z odpowiedniej bazy danych, w momencie wybrania konkretnego produktu i specyfikacji. Wszystko, co powinien wiedzieć użytkownik, to jak nazwana została się analiza, którą chce wykonać.

Właśnie tak działa system **SEWSS!**

Wprowadzenie do **SEWSS**

SEWSS to wieloużytkownikowe oprogramowanie umożliwiające korzystanie z wielu różnorodnych funkcji. Najważniejsze elementy systemu **SEWSS** to:

- **Połączenia ze źródłami danych.** **SEWSS** konfigurowany jest tak, aby mógł korzystać z repozytoriów danych w celu analizy i wizualizacji danych. Każdemu szablonowi odpowiada źródło danych, a wymagane dane pobierane są automatycznie.
- **Szablony analizy.** W systemie **SEWSS** szablony (wzorcy) analiz są przechowywane centralnie i mogą być uruchamiane automatycznie lub na żądanie użytkownika. Szablony analizy mogą mieć ustalone ustawione wszystkie parametry, albo użytkownik może część z nich określać przy uruchamianiu analizy (np. wybierając podzbiór analizowanych danych lub konkretną wersję analizy).
- **Interakcyjne filtrowanie:** **SEWSS** zawiera niezawodne narzędzia filtrowania danych. Dzięki temu użytkownik nie musi poznawać złożonych struktur danych i języka SQL. Zamiast tego, użytkownik wybiera tylko odpowiednie wartości (np. nazwę produktu, zakres daty i czasu) i zadaje kryteria dla danych.

- **Raporty.** System **SEWSS** umożliwia korzystanie z modyfikowalnych szablonów raportów, w których możemy określić układ i format strony dla regularnie tworzonego raportu. W skład **SEWSS** wchodzi aplikacja automatycznie tworząca raporty, w zadanych odstępach czasu.
- **Bezpieczeństwo i uprawnienia.** W systemie wieloużytkownikowym nie powinno się każdego użytkownika traktować tak samo. Zadania i zakres odpowiedzialności są różne dla różnych osób. W systemie **SEWSS** zapisywane są uprawnienia do przeglądania i modyfikowania szablonów analiz oraz do korzystania z funkcji systemu. Łatwość korzystania z systemu zwiększa to, że użytkownik nie widzi obiektów, które nie są mu potrzebne.

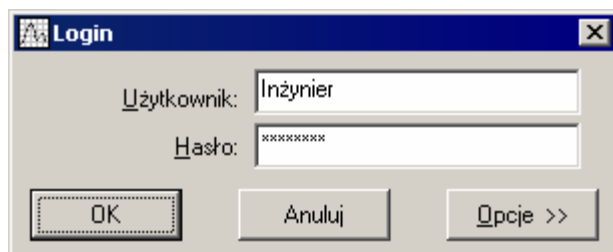
Dzięki powyższym funkcjom system **SEWSS** dodaje do repozytoriów danych praktyczne możliwości analizy. **SEWSS** jest gotowym produktem i zawiera ogólne obiekty, które są konfigurowane za pomocą wbudowanego w system interfejsu administratora. Najważniejsze obiekty systemu opisano poniżej. Zapoznanie się z podstawowymi terminami **SEWSS** daje podstawę do zrozumienia przykładów przedstawionych niżej.

- **Połączenia.** W połączeniach zawarte są określenia źródeł danych i metod dostępu do nich.
- **Profile.** Profile to zbiory zapytań, które określają dane podlegające analizie. W profilach zapisywane są również meta-dane, umożliwiające odpowiednią interpretację danych podczas analizy (np. to, że w pewnej kolumnie zapisano oznaczenie czasu, a w innej nazwę produkty, względem której filtrujemy dane).
- **Monitory.** Monitory to szablony analiz, określające rodzaj analizy i jej parametry.
- **STATISTICA Visual Basic:** **SEWSS** jest zorientowaną obiektowo aplikacją wyposażoną w interfejs API realizowany przez język Visual Basic for Applications (VBA). Działaniem **SEWSS** możemy sterować nie tylko interakcyjnie, ale również za pośrednictwem ponad 11 000 funkcji modelu obiektowego **STATISTICA**.

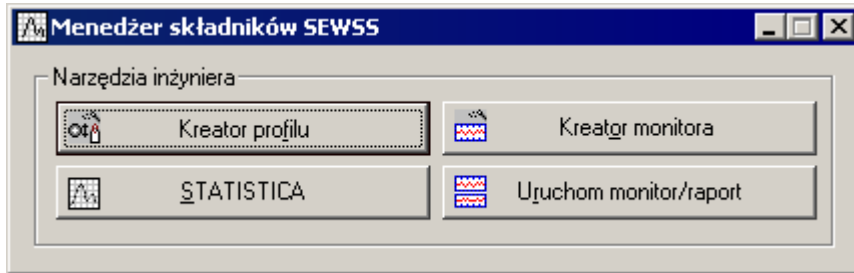
Przykład: Współpraca z bazą danych

W niniejszym rozdziale przedstawiamy przykład wykorzystania systemu **SEWSS** jako narzędzia analitycznego połączonego z repozytorium danych. Zdefiniujemy połączenie ze źródłem danych i profil określający, które dane będą analizowane. Na koniec zdefiniujemy same analizy.

Pracę musimy zacząć od zalogowania się w systemie. Podany przez użytkownika identyfikator określa, co będzie mógł on zobaczyć i jakie działania wykonywać.



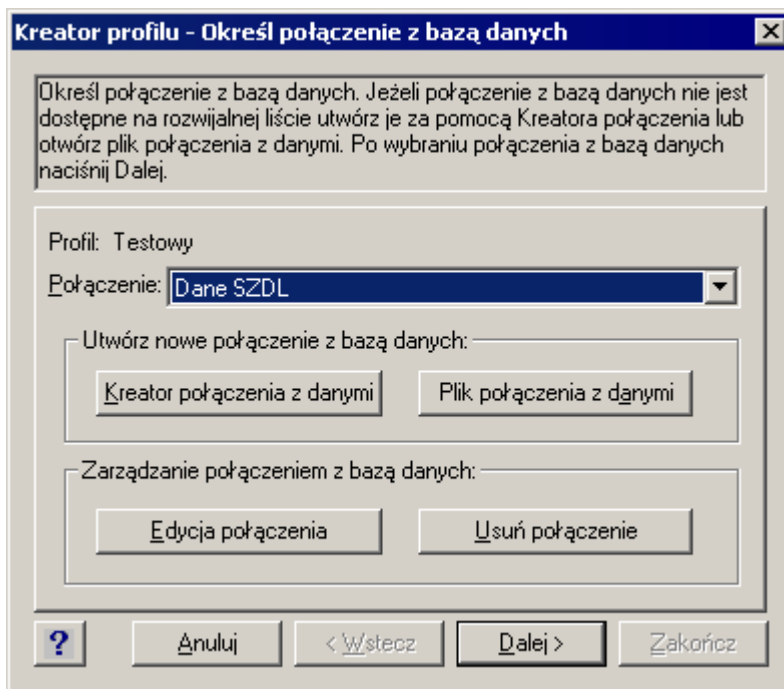
Po sprawdzeniu hasła i nazwy użytkownika, system wyświetla główny interfejs użytkownika, zwany *Menadżerem składników*. Z *Menadżera składników* uruchamiamy aplikacje systemu, przy czym zestaw dostępnych aplikacji zależy od uprawnień użytkownika.



W tym przypadku użytkownik jest uprawniony do definiowania profili i monitorów. Wykonywanie tych czynności ułatwiają kreatory.

Warto zwrócić uwagę, że większość użytkowników **SEWSS** to odbiorcy informacji, korzystający z wcześniej zdefiniowanych zapytań, szablonów analiz i raportów. Konfigurowanie systemu wykonywane jest przez niewielu użytkowników i tylko wtedy, gdy określamy nową analizę lub gdy zaszły jakieś zmiany.

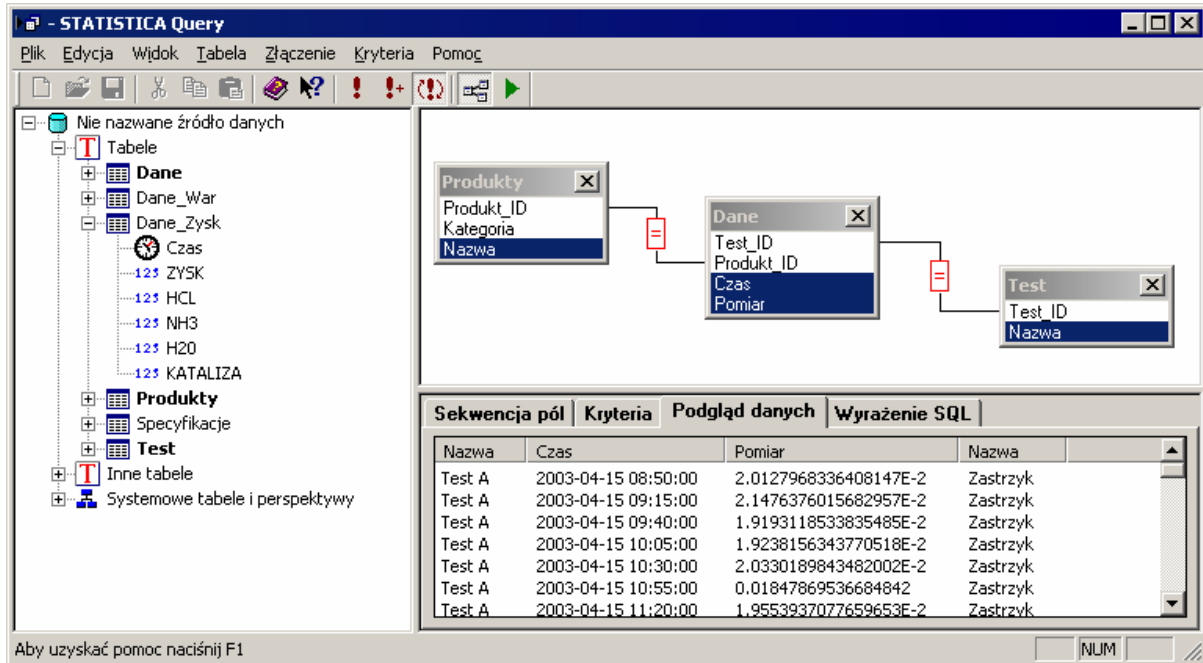
W tym przykładzie będziemy pobierać dane z bazy danych „SZDL” (System Zarządzania Danymi Laboratoryjnymi). Najpierw należy zdefiniować połączenie do tej bazy danych.



Po zdefiniowaniu połączenia, możemy je wykorzystać w jednym lub wielu profilach pobierających dane. Dostępne połączenia znajdują się na liście rozwijalnej *Połączenie* (zob.

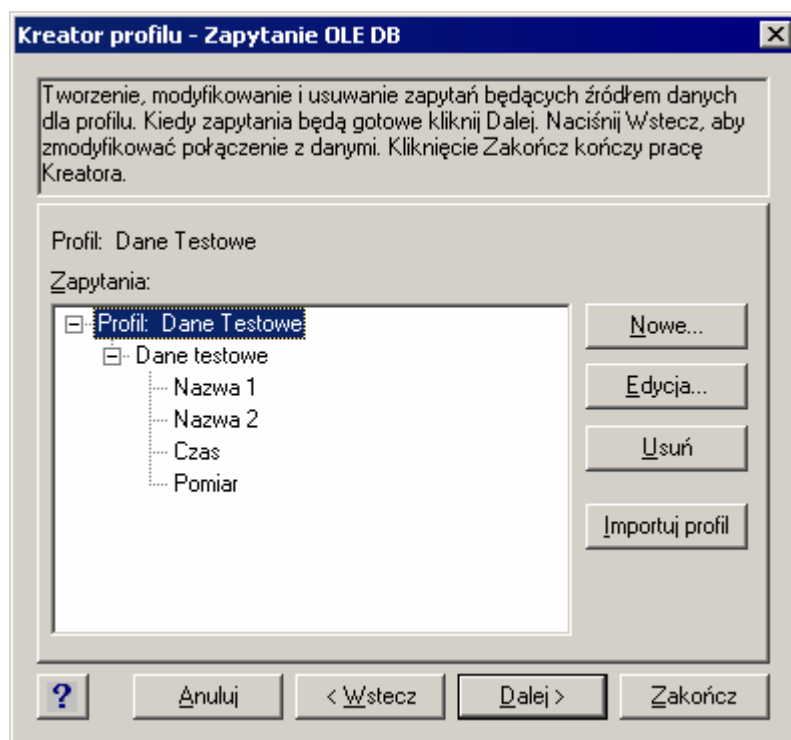
rysunek powyżej), z której korzystać mogą osoby o uprawnieniach administratora systemu **SEWSS**.

Profil to zestaw zapytań określających analizowane zbiory danych. W systemie **SEWSS** zapytania budujemy w przyjaznym użytkownikowi środowisku graficznym, pokazanym na poniższym rysunku:



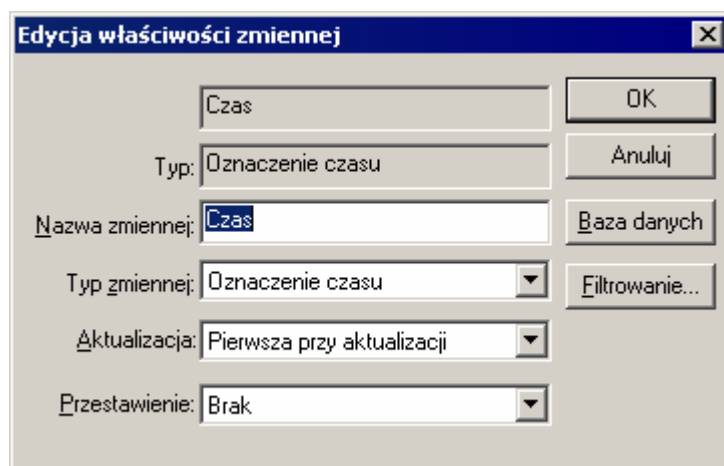
Podczas budowy zapytania, na karcie *Podgląd danych* możemy sprawdzić, czy zapytanie zwraca odpowiednie dane.

Zestaw zapytań wchodzących w skład profilu, jest widoczny w oknie kreatora profilu (zob. rysunek poniżej).



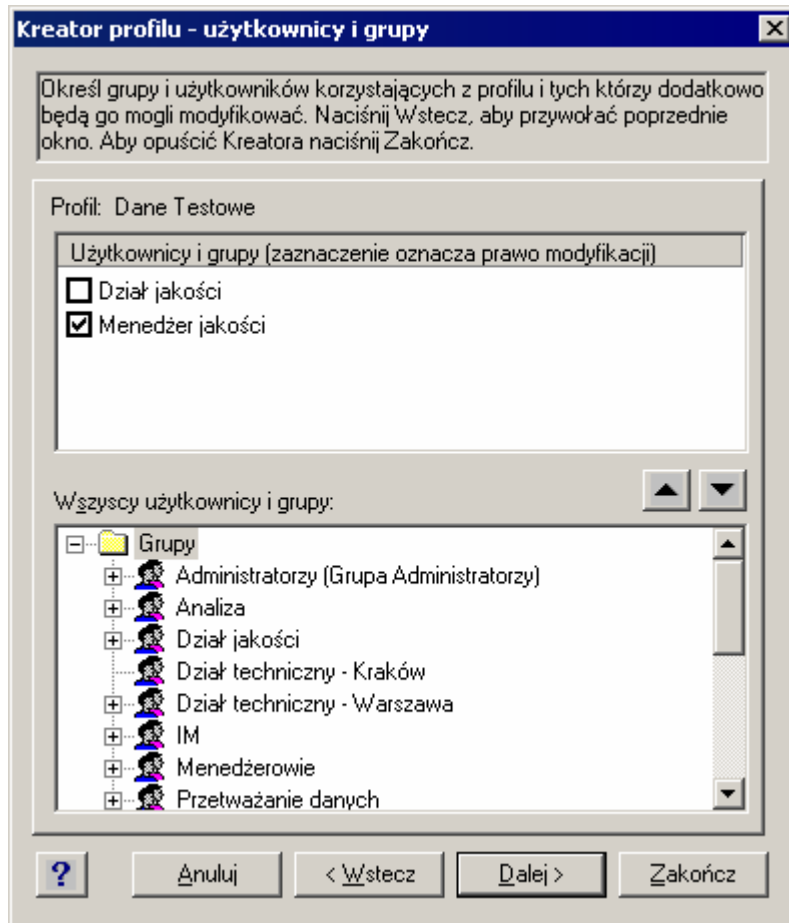
W naszym przypadku zdefiniowaliśmy jedno zapytanie o nazwie „Dane testowe”. Zapytanie to zwraca dane potrzebne do analizy kilku testów produktów, jest tu więc nazwa produktu, nazwa testu, czas pomiaru i jego wynik.

W następnym kroku administrator określa metadane dla zapytania. Metadane wskazują jak powinny być traktowane wyniki zapytania. Przykładowo, na poniższym rysunku widzimy, że kolumna *Czas* będzie traktowana jako oznaczenie czasu, a jej wartości wykorzystywane będą przy automatycznej aktualizacji danych do rozpoznawania nowych danych, które pojawiły się po ostatnim wykonaniu zapytania przez **SEWSS**.



Dostęp do danych w systemie **SEWSS** może być ograniczany dzięki możliwości przypisania użytkownikom odpowiednich uprawnień. Podczas definiowania profilu, administrator

wskazuje, którzy użytkownicy będą mogli z niego korzystać; tak jak pokazano to na poniższym rysunku:



Po określeniu profilu i zapisaniu go w systemie **SEWSS**, można go użyć jako podstawy jednego lub wielu szablonów analiz. Innymi słowy, jedną z najważniejszych cech systemu, jest możliwość wykonywania dla tych samych danych wielu różnych analiz. Takie podejście jest powszechne choćby dlatego, że różne osoby w zależności od swoich funkcji, zainteresowane są różnymi aspektami danych.

Definiując monitor określić trzeba profil, na którym będzie on oparty. Dla ułatwienia można założyć, że monitor będzie dostępny dla tych samych użytkowników i grup co profil, oraz że będzie dołączony do tego samego węzła.

Kreator monitora - Cechy monitora

Podaj nazwę, opis, profil i typ dla monitora, a następnie naciśnij Dalej.
Naciśnij Wstecz aby powrócić do poprzedniego okna. Aby zakończyć pracę Kreatora kliknij Zakończ.

Nazwa:
(Pozostaw pustą, aby użyć nazwy profilu)

Opis:

Profil:

Typ:

Umieść w Schemacie systemu tam, gdzie profil
 Użytkownicy i grupy tacy jak dla profilu

Podczas definiowania monitora tworzącego kartę kontrolną można, między innymi, określić z jakiego okresu czasu dane mają pojawić się na karcie.

Kreator monitora - Definiowanie danych

Zapytania | Kryteria | Data i czas | Opcje

Zastosuj do
 Wszystkich zapytań Zapytania:

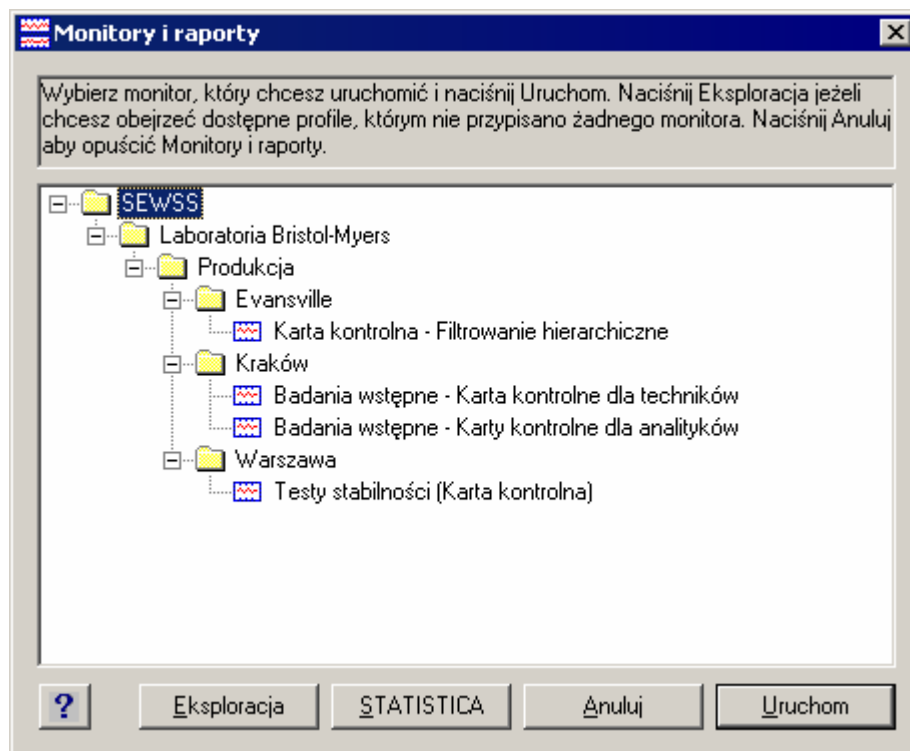
Przedział czasu
Od:
Do:

Określenie daty i czasu
 Bez danych początkowych
 Podana przez użytkownika
 Dziś
 Bieżący tydzień
 Bieżący miesiąc
 Bieżący kwartał
 Dane z ostatnich

Pobierz wszystkie dane

Oprócz zakresu danych wykreślanych na karcie istnieje także możliwość określenia statystyk, jakie pojawiać się mają na kartach kontrolnych, dla poszczególnych właściwości, a także zdefiniowania automatycznych reakcji systemu na rozregulowania procesów.

Po zdefiniowaniu monitora, użytkownicy analiz o odpowiednich uprawnieniach, zobaczą go w *Schemacie systemu*, tak jak pokazano to na poniższym rysunku:



Podsumowanie

SEWSS jest potężną, elastyczną i łatwą w użyciu platformą dla potrzeb data mining, analizy danych i prezentacji wyników dla całości organizacji. Materiał ten zawierał wstępne przedstawienie cech i zalet **SEWSS**, na przykładzie jego zdolności do współpracy z bazami danych.

W celu bardziej szczegółowego zapoznania się z możliwościami **SEWSS** lub urzędowania indywidualnej prezentacji systemu, dostosowanej do konkretnych potrzeb, prosimy o kontakt ze StatSoft Polska (tel. 012-4284300; e-mail: info@statsoft.pl).

Niniejszy dokument: "Część 1: Współpraca z bazami danych", jest pierwszym z grupy dokumentów prezentujących możliwości **SEWSS**, dotyczące współpracy z bazami danych. Kolejnych informacji dostarczają teksty: "Część 2: Interaktywne filtrowanie danych" oraz "Część 3: Przypisywanie specyfikacji".