

Przegląd ciekawych elementów SQL 2008

Tomasz Kopacz
tkopacz@microsoft.com

Plan

Uwaga – wersja CTP/Beta

- Ogólne możliwości SQL Server 2008
- Motor relacyjny i TSQL
- Różne czynniki związane z wydajnością
- Nierelacyjne typy danych
- Kilka uwag o dostępie do danych
- Niestety pominięte:
 - Zarządzanie; policy itp.
 - Analiza i eksploracja danych (wymaga oddzielnej sesji)

Prezentacja – referencja; dużo slajdów ukrytych – czas...

Podstawowe składniki SQL 2008

Motor relacyjny

- Też: Spatial, Hierarchie, CTE (rekurencje), XML, FILESTREAM, FullText, programowanie w .NET....

ETL, integracja

- SQL Server Integration Services
- Linked Server i okolice

Business Intelligence

- Czyli: OLAP, Data Mining i przeglądarki

Reporting Services

- Designer, ad hoc Report Modeler

Zarządzanie

- Dużo...

Relacyjny motor danych

(Wybrane) rozszerzenia T-SQL

- CREATE/ALTER SCHEMA
- Synonimy
- nvarchar(max)
- SELECT TOP(@c)
- TABLESAMPLE
- EXCEPT/INTERSECT
- INSERT/UPDATE ...
OUTPUT
- CREATE EVENT
NOTIFICATION
- ROW_NUMBER()/RANK()
- Triggery na DDL
- CROSS
APPLY/OUTER APPLY
- TRY / CATCH
- Catalog views
- DECLARE
@variable int = 2;
- INSERT INTO table
(colA, colB, colC)
VALUES (1, 2, 3), (4, 5,
6);

Nowe typy danych

Date

- 3 bajty, tylko data
- Od 0001-01-01 do 9999-01-01 kalendarzu gregoriański

Time (n)

- 3-5 bajtów, tylko czas
- Dane z precyzją do 100 ns

DateTimeOffset(n)

- 8-10 bajtów, strefy czasowe / czas UTC
- Dane z precyzją do 100 ns

DateTime2 (n)

- 6-8 bajtów, 0001-01-01 do 9999-01-01
- Bez uwzględniania stref czasowych

DateTime

- (dla porównania) 8 bajtów; dokładność 0.33 sekundy

Smalldatetime

- (dla porównania): 4 bajty; dokładność 1 minuta

Rowversion

- Dawny TimeStamp; pewne zmiany

Dodatkowe funkcje:

`SYSDATETIME/SYSDATETIMEOFFSET/SYSUTCD
ATETIME/SWITCHOFFSET/TODATETIMEOFFSET`

Przykład(y) DateTime

```
CREATE TABLE t1 (c1 DATE, c2 TIME(3),
  c3 DATETIME2(7) NOT NULL DEFAULT GETDATE(),
  c4 DATETIMEOFFSET CHECK
    (c4<CAST(GETDATE() AS DATETIMEOFFSET(0)))
);
```

```
INSERT INTO t1 VALUES ('0001-01-01', '23:59:59',
  '0001-12-21 23:59:59.1234567',
  '0001-10-21 23:59:59.1234567 -07:00');
```

```
INSERT INTO t1 VALUES ('9999-12-31', '23:59:59',
  '9999-12-31 23:59:59.1234567',
  '1111-10-21 23:59:59.1234567 -07:00');
```

```
SELECT c4,
  DATEPART(TZOFFSET, c4),
  DATEPART(ISO_WEEK, c4),
  DATEPART(MICROSECOND, c4) FROM t1;
```

Rozszerzony typ tabelaryczny

- Zmienna - „Zbiór”
- Wygodniejsze niż tabela tymczasowa
- Naturalny czas życia
 - Zakres: SP/funkcja itp
- Można definiować indeksy / więzy
- Sys.table_types
- Parametr wejściowy dla SP
- De facto - BCP wewnątrz serwera
- ADO.NET
 - SqlDbType.Structured

```
create type t_tvptable as
table(i int)

create proc usp_tbpinsert
(@p t_tvptable readonly) as
    insert t select * from @p
```


Merge

- Wybór operacji w zależności czy warunek spełniony czy nie
- Scenariusz: aktualizacja magazynu
 - Dodanie nowego wiersza gdy towar nie istnieje
 - Aktualizacja ilości gdy towar istnieje

```
MERGE INTO dbo.Quantities Q
USING dbo.Transactions T ON T.ProductID = Q.ProductID
WHEN MATCHED THEN
    UPDATE SET Q.Quantity = Q.Quantity + T.Quantity
WHEN NOT MATCHED THEN
    INSERT (ProductID, Quantity) VALUES (T.ProductID, T.Quantity);
```

GROUPING SETS

- SQL 2005 – tabela przestawna
 - PIVOT
 - UNPIVOT
- SQL 2008 grupowanie w zdefiniowanych zbiorach

	City	Country	Continent	Planet
1	Berlin	Germany	Europe	Earth
2	Warsaw	Poland	Europe	Earth
3	Lublin	Poland	Europe	Earth
4	Toronto	Canada	North America	Earth
5	New York	USA	North America	Earth
6	Cordoba	Argentina	South America	Earth

```
SELECT Planet, Continent, Country, Count(*)  
as CitiesCount  
FROM Cities  
GROUP BY  
GROUPING SETS (Planet, Continent, Country)
```

	Planet	Continent	Country	CitiesCount
1	NULL	NULL	Argentina	1
2	NULL	NULL	Canada	1
3	NULL	NULL	Germany	1
4	NULL	NULL	Poland	2
5	NULL	NULL	USA	1
6	NULL	Europe	NULL	3
7	NULL	North America	NULL	2
8	NULL	South America	NULL	1
9	Earth	NULL	NULL	6

Funkcja GROUPING_ID

Drobiazgi związane z SQL

Demo

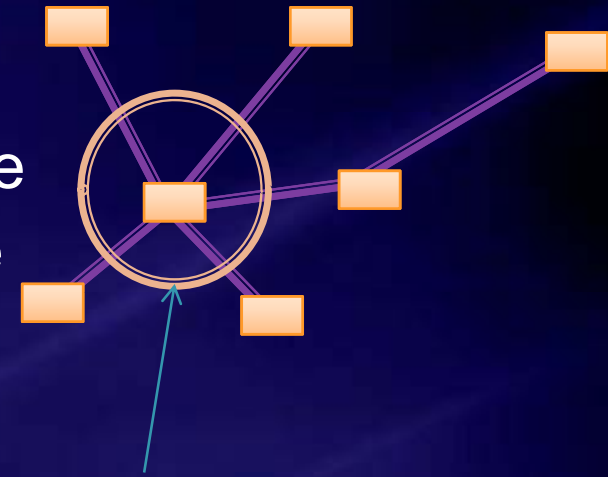
Różne czynniki związane z wydajnością

Wydajność SQL 2008

- Kompresja danych (i kopii zapasowej)
- Tryby optymalizacji kwerend
 - „tuning mode” -> cache, wybór najlepszego
- Hint Forceseek
- „Przewidywalna” wydajność
- Resource Governor
 - Per „typ” kwerendy, aplikacja, użytkownik...
- Zarządzanie i analiza wydajności
 - Performance Data Collection:
 - Wszystkie dane istotne z punktu widzenia wydajności
 - Performance Data Warehouse
 - Performance Studio

Problem z zapytaniem Star Join

- Używane w raportach/analizach OLAP
- Grupowanie / agregacja zwalnia operacje
- Z tabeli faktów niewiele jest eliminowane w początkowych etapach optymalizacji
- Nawet jeden wiersz w wyfiltrowanym wymiarze „generuje” dużą liczbę pasujących wierszy w tabeli faktów.
- Rozwiązanie:
 - Rozpoznać wzorzec
 - Statystyki, sposób sformułowania tabeli
 - Dostępne od SQL Server 7.0
- Alternatywny plan kwerend dla Star Join
 - Zmiana sposobu liczenia kosztu - dokładniejszy
- Filtr bitmapowy



**Dodatkowy komponent:
Patter Detection**

Plany wykonania kwerend

- (procedura składowana do ustalenia planu dla danego wyrażenia SQL)
- SSMS / SMO (cywilizowane UI)
- Postępowanie:
 - Tuning aplikacji
 - Krytyczne kwerendy do pamięci
 - „Zamrozić” pamięć `sp_create_plan_guide_from_cache`
 - Optymalizację robi SQL – ew. hinty, ale potem – i tak zapisany PLAN
 - Query Profiler generuje plany itp.
- Walidacja planu

- System Databases
 - master
 - model
 - msdb
 - tempdb
- Database Snapshots
- AdventureWorks
 - Database Diagrams
 - Tables
 - Views
 - Synonyms
 - Programmability
 - Stored Procedures
 - Functions
 - Database Triggers
 - Assemblies
 - Types
 - Rules
 - Defaults
 - Plan Guides
 - Guide1_All_Stateme
 - Guide1_All_Stateme
 - Guide1_Statement2
 - PlanGuide-20070717
 - PlanGuide-20070717
 - PlanGuide-20070717
 - Service Broker
 - Storage
 - Security
 - MDW
- Security
- Server Objects
- Replication
- Management
- SQL Server Agent

Plan Guide Properties

Ready

Hints	<ShowPlanXML xmlns="http://schemas.microsoft.com/sqlserver/2004/07/showplan" Version="1.0" Build="10.0.1051.2"><E
IsDisabled	False
Name	Guide1_All_Statements_998
Parameters	
ScopeBatch	SELECT * FROM Production.Product WHERE ProductSubcategoryID > 4; □ □ SELECT * FROM Person.Address; □ □
ScopeObjectName	
ScopeSchemaName	
ScopeType	Sql
Statement	SELECT * FROM Person.Address; □ □

Name

OK Cancel Help

- [-] Programmability
 - [+] Stored Procedures
 - [+] Functions
 - [+] Database Triggers
 - [+] Assemblies
 - [+] Types
 - [+] Rules
 - [+] Defaults
 - [-] Plan Guides
 - Guide1_All_States
 - Guide1_States
 - PlanGuide-
 - PlanGuide-
 - PlanGuide-
- [+] Service Broker
- [+] Storage
- [+] Security
- [+] Security
- [+] Server Objects
- [+] Replication
- [+] Management
- SQL Server Agent (Agent)

- New Plan Guide...
- Script Plan Guide as
- Enable
- Disable
- View Dependencies
- Reports
- Delete
- Refresh
- Properties

```

CROSS APPLY sys.dm
WHERE st.text LIKE

print @xml_showplan
EXEC sp_create_plan_guide
    @name = N'Guide1_from_XML_showplan',
    @stmt = N'SELECT City, StateProvinceID
    ,
    @type = N'SQL',
    @module_or_batch = NULL,
    @params = NULL,
    @xml_showplan;

sys.plan_guides

cache

StateProvinceID, PostalCode F

EXECUTE To

```

Scenariusze użycia

- Łatwa aktualizacja
- Blokada zmian planu dla krytycznych kwerend/planów
- Łatwe rozwiązywanie problemów z kwerendami
 - Bez hintów!
- Można „naprawić” złą aplikację
- Takie samo zachowanie serwera testowego i produkcyjnego

Inne

- Rzadkie kolumny; dużo kolumn
 - COLUMN_SET FOR ALL_SPARSE_COLUMNS...
 - Przechowuje „wskaźnik” + wartość: (kolumna1, kolumna100) (1,9)
- Filtrowane indeksy
 - Utwórz indeks dla A=50 (A<>NULL itp.)
- Change Data Capture (I np. zasilenie hurtowni)
- Znacznik czasu w „wierszu” (Pokaż wiersze zmienione w...)
- Vardecimal: (((sign) exponent) (mantissa))
 - sys.sp_estimated_rowsize_reduction_for_vardecimal
 - sp_tableoption '<table>', 'vardecimal storage format', 1
 - Opłacalne _czasami_ (do mniej więcej 30 cyfr znaczących)
- Kompresja
 - Wiersz o zmiennej długości
 - Kompresować można ROW i/lub PAGE
 - Jak: atrybut DATA_COMPRESSION przy DML; np. dla partycji
 - sp_estimate_data_compression_savings

FORCESEEK, CDC, ...

Demo

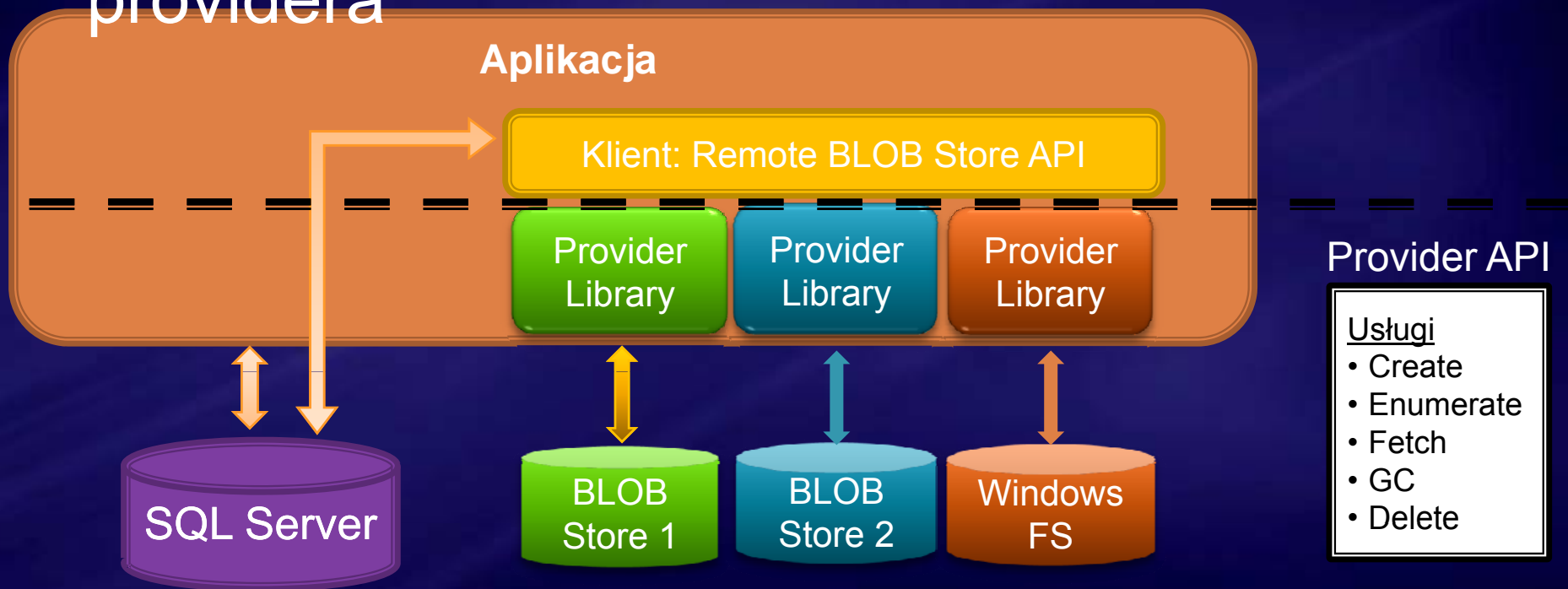
Nierelacyjne typy danych

Możliwości

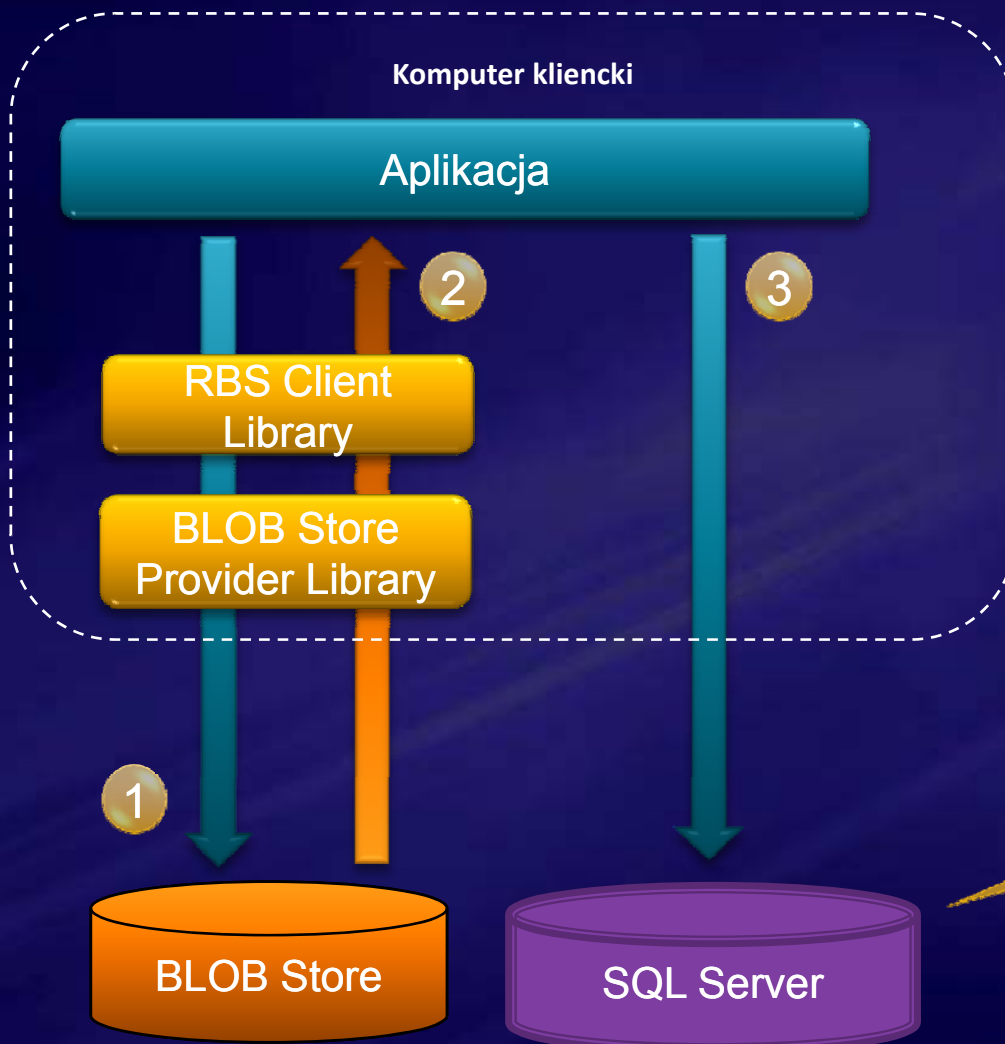
- Przechowywanie dużych obiektów binarnych
 - Plik + link w bazie
 - BLOB (varbinary (max) itp.)
 - Zdalny BLOB
 - FILESTREAM
 - Reinkarnacja WinFS
- XML
- Zintegrowany Full Text Search
- Duże typy użytkownika w .NET

Remote BLOB Store API

- Aplikacja w sposób przeźroczysty używa dowolnych pojemników BLOB
- Producent pojemnika BLOB dostarcza providera



Użycie „zdalnego” BLOB

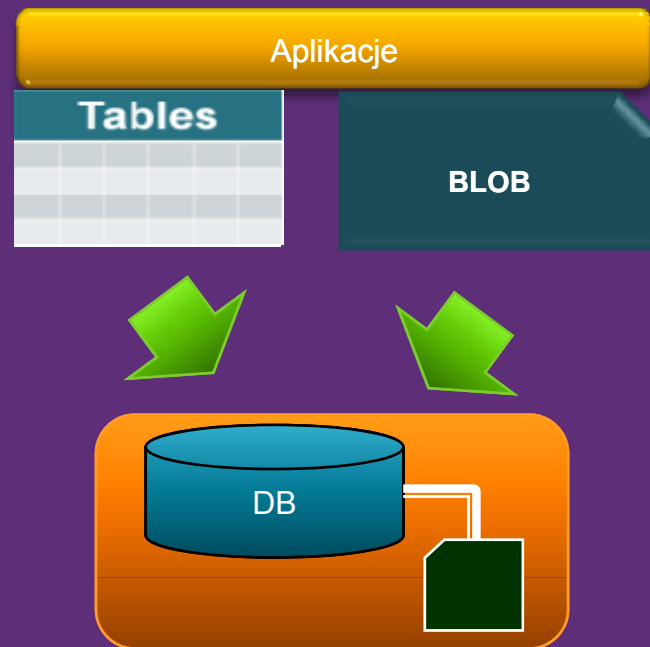


- 1 Zapis BLOB (np. zdjęcie)
- 2 Zwraca BlobReference
- 3 Zapis BlobReference do PhotoRef field

ClaimID	ClaimDate	PhotoRef
4390	6/5/2007	<xml>

Filestream

BLOBs w DB + system plików



- Plik w NTFS + link w bazie
 - Zarządzane przez SQL-a
- De facto atrybut na polu VARBINARY(MAX)
- Dwa modele programowania
 - TSQL (znany SQL BLOB)
 - Win32 Streaming APIs z semantyką transakcji
- Gwarancja spójności danych
- Wspólne zarządzanie
 - Kopia zapasowa
 - Uprawnienia SQL-owe
- Limit rozmiaru: wielkość dysku(ów)

Zarządzanie hierarchiami

- Scenariusze użycia
 - Tematy na forach, listach itp
 - Diagram zależności służbowej
 - Kategorie (produktu, CMS itp.)
 - System plików
 - Geneologia
- SQL 2005: CTE (Common Table Expression)
 - Rekurencja w operatorach na zbiorach

HierarchyId

- Wbudowane metody do manipulacji hierarchiami
 - Na przykład: GetAncestor , GetDescendant , GetLevel
 - INSERT #Hierarchy VALUES ('Debbie Walker','/1/2/')
- Indeksy depth-first i breadth-first
- Nadal jest to baza RELACYJNA



XML

- 2 zastosowania:
 - Mapowanie XML <-> relacyjne
 - Baza „hierarchiczna” w bazie relacyjnej
- Klient:
 - SqlXml (System.Data.SqlTypes)
 - XmlReader z `sqlxml.CreateReader()`
 - DataSet – typ pola `XPathDocument`
 - Usługa Web itp.
- Typ kolumny/zmiennej
 - Dokument XML 1.0 lub fragment
 - Przechowywane:
 - Jako LOB (2 GB), binarnie
 - Zakodowane w UTF-16
 - W3CSchema / bez schematu
- Ładnie się łączy z SQL!
 - Też SP: `sql:variable("@xml")`
- CAST/CONVERT
- Indeksy
 - `CREATE PRIMARY XML INDEX idx_1 ON docs (xDoc)`
 - Dodatkowe indeksy na wartościach, (VALUE), ścieżkach (PATH), właściwościach (PROPERTY)
- Lax validation (anyType), `xs:DateTime...`
- XQuery
 - `query()`, `value()`, `exist()`
 - `SELECT id, xDoc.query ('for $s in /doc[@id = 123]//section...`
`FROM docs`
 - `let $count :=count($invoice/Items/Item)`
 - `modify()` (XML DML)
 - `UPDATE docs SET xDoc.modify('insert <section num="2">...`
`after (/doc/section[@num=1])[1]')`

Dane przestrzenne (Spatial)

- Typy: Geometry i Geography
- Indeksy Spatial
- Standardowa składnia SQL; pełny zestaw składników Open Geospatial Consortium (OGC/SQL MM, ISO 19125)
 - Typy :geometry, geography
 - Metody: intersects, buffer, itp
 - Indeksy przestrzenne
- Integracja z Virtual Earth

Zmiany w integracji w CLR

- SQL 2005: typy, metody, procedury składowane, agregaty itp.
 - Indeksowane pod pewnymi warunkami
- SQL 2008:
 - Wsparcie dla typów nullable (int?)
 - Obsługa dużych instancji (>8000)
 - Zmiany w agregatach
 - Duże wyniki pośrednie (>8000)
 - Wejście wielokolumnowe
 - Funkcje zwracające tabele
 - Można określić porządek sortowania
 - Używa optymalizator!

```
CREATE FUNCTION myTVF()  
RETURNS TABLE (@i int, @j int, @k int)  
WITH ORDER i ASC, j DESC  
AS EXTERNAL NAME ...
```

Hierarchie, XML, ...

Demo

Audyt, szyfrowanie i okolice

Audyt

- Cel: śledzenie aktywności w bazie na poziomie biznesowym
- Obiekt AUDIT
- Poziom ziarnistości:
 - Baza, obiekt (tabela, kolumna, użytkownik)
- Wiele wyjść (plik, log Windows)
 - Konsolidacja – Analysis Services, Reporting Services itp.

Audyt - konfiguracja

```
CREATE AUDIT HIPAA_Audit  
TO FILE  
(  
  FILENAME='\\PRO1\Aud\HIP_ADT.aud',  
  MAX_SIZE=100 MB,  
  RESERVE_DISK_SPACE  
)  
WITH (SHUTDOWN_ON_FAILURE = ON);
```

```
CREATE AUDIT SPECIFICATION  
SvrAC  
ON SERVER  
TO HIPAA_Audit  
  ADD FAILED_LOGIN_GROUP;
```

```
CREATE AUDIT SPECIFICATION  
AuditAC  
ON DATABASE  
TO HIPAA_Audit  
  ADD SELECT ON  
  table::Customers(payment);
```

Extended Events (xEvent)

- Podstawa audytu w SQL 2008
- Lekki i wydajny infrastruktura do śledzenia
- Łatwa instrumentalizacja
- Zintegrowana z ETW (Event Tracing for Windows)
 - Można skorelować różne zdarzenia
- Wykorzystywane przez przyszłe narzędzia do automatycznej administracji produkcyjnej i analiz „problemów”

Przeźroczyste szyfrowanie danych

- Można szyfrować na poziomie kolumn
 - Klucz związany z loginem/grupą
 - Specjalna składnia poleceń
- SQL 2008 – zakres bazy danych
 - Database Encryption Key (DEK)
 - Też kopie zapasowe
- Po prostu:
 - CREATE DATABASE ENCRYPTION KEY
 - WITH ALGORITHM = AES_128
 - ENCRYPTION BY SERVER
 - CERTIFICATE ent_cert

Zarządzanie kluczami

- Integracja z innymi systemami zarządzania kluczami
 - Współdzielenie, zarządzanie wiekiem, dedykowany sprzęt
- Współpraca z zewnętrznymi providerami: RSA, Ncipher, Safenet

Analiza danych

UDM, OLAP, SSIS

- Unified Data Model

- Pozwala nazywać kolumny, tłumaczyć...
- Dodać relacje których nie ma w bazie
- Automatycznie generuje (przy zachowaniu pewnych konwencji)

- SQL Server Integration Services; SSIS (dawniej DTS)

- Dwa główne elementy – dane i przepływ
- Rozbudowane środowisko graficzne
- Łatwo rozszerzalne (naprawdę łatwo)
- Parametryzowane „z zewnątrz” IDE; skrypty i okolice

- Business Intelligence Studio

- Tryb on/offline
- Ostrzeżenia projektowe, generatory (konwencja nazewnicza)
- (kostki, wymiary, raporty, datamining)
 - Praca zespołowa (TFS)

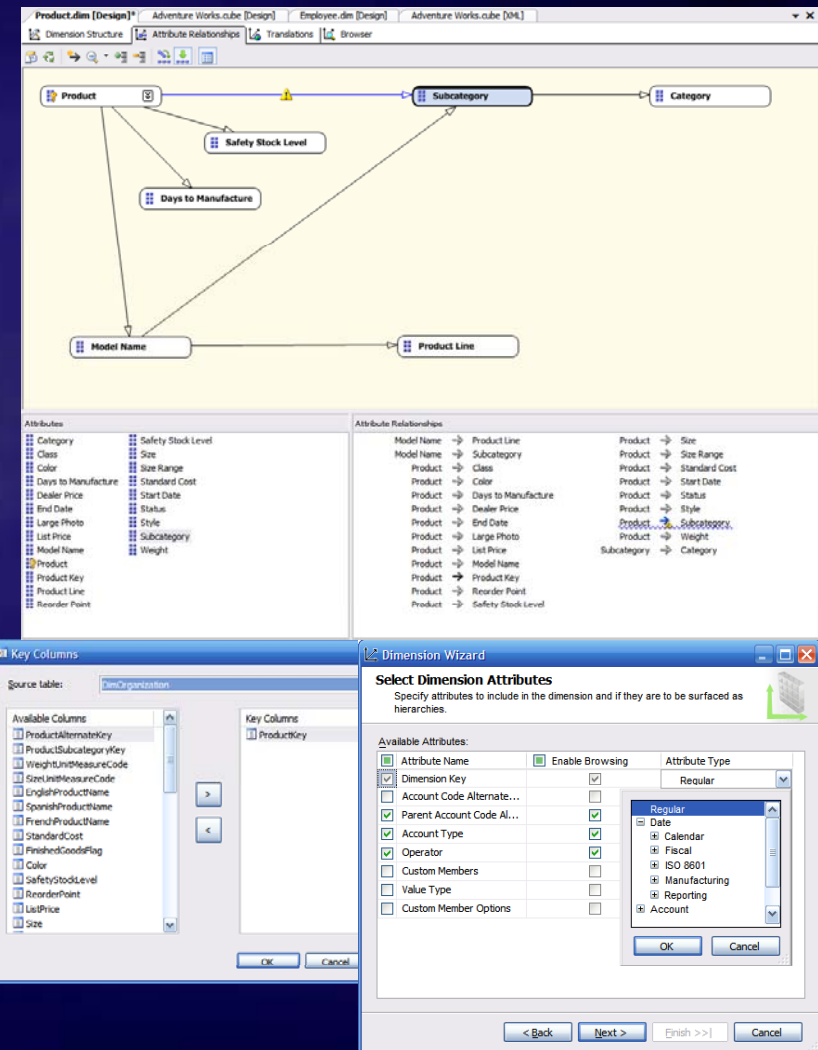
- OLAP

- Kilka tabel faktów; Osierocoone wiersze do kategorii „nieznane”
- Tłumaczenie terminów na różne języki
- Perspektywa na kostce
- Web services + Natywny XML/A
- Pro-active caching (MOLAP -> ROLAP, ale działa szybko)
- Wymiary – bez ograniczenia ilości elementów
 - Referencyjne; Do tabeli faktów; Many-to-many; kilka hierarchii; własne grupowanie (dowolny układ geograficzny; okresy rozrachunkowe)
- KPI
- MDX
 - Object Definition Language
 - Specjalne wyliczanie z pomijaniem „pustych” elementów

Kilka zrzutów...

Projektowanie wymiarów

- **Attribute Relationship Designer**
 - Łatwa edycja związków pomiędzy atrybut
 - Najważniejsze – walidacje i sugestie dobrego (optymalnego) projektu
- **Dimension Wizard**
 - Automatyczne parent-child
 - Składniki „klasyfikujące”
 - Błędy i ostrzeżenia/rady
- **Dimension Editor**
 - Łatwiejsze definiowanie kluczy
 - Wygodna siatka (Grid) 😊



Projektant Agregacji

Notes
View attribute relationship changes when aggregation is changed

View a single aggregation – Multiple panes

Partition Name

Aggregation Viewer Create New Aggregation Sort Aggregations Select Aggregation Search

Database: name of database

Cube: name of cube

Measure Group: name of measure group

Select Aggregation Design

Controls

View Dimensions Expand All Collapse All more...

View Attributes All Top resource consuming Most Used Only Selected Only Deselected

Preferences Saved Views Import Aggregation? More..

Related Dimensions		Edit Aggregation				
		Aggregation (001)	Aggregation (002)	Aggregation (003)	Aggregation (004)	Aggregation (005)
		Estimated Rows 00	Estimated Rows 00	Estimated Rows 00	Estimated Rows 00	Estimated Rows 00
Dimension (A)	Attribute (A1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (A2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (A3) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (A4) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dimension (B)	5 Attributes	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (C1) Count 000	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Dimension (C)	Attribute (C2) Count 000	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (C3) Count 000	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (C4) Count 000	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (D1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Attribute Relationships (for selected Dimension) | Dimension (C)

Zoom Control Control Control

View all aggregations – Edit Aggregation

Partition Name

Aggregation Viewer Create New Aggregation Sort Aggregations Find Aggregation Search

Database: name of database

Cube: name of cube

Measure Group: name of measure group

Select Aggregation Design

Controls

View Dimensions Expand All Collapse All more...

View Attributes All Top resource consuming Most Used Only Selected Only Deselected

Preferences Saved Views Import Aggregation? More..

		Edit Aggregation				
		Aggregation (001)	Aggregation (002)	Aggregation (003)	Aggregation (004)	Aggregation (005)
		Estimated Rows 00	Estimated Rows 00	Estimated Rows 00	Estimated Rows 00	Estimated Rows 00
Dimension (A)	Attribute (A1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (A2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (A3) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (A4) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dimension (B)	Attribute (B1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (B2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (B3) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (B4) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dimension (C)	Attribute (C1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (C2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (C3) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (C4) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dimension (D)	Attribute (D1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (D2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (D3) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (D4) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dimension (E)	Attribute (E1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (E2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Attribute (E3) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (E4) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dimension (F)	Attribute (F1) Count 000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Attribute (F2) Count 000	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Replace Existing Save As Cancel Save As

Data Mining

- Algorytmy podstawowe
 - Decision Trees, Clustering, Naïve Bayes, Time Series, Association, Sequence Clustering, Neural Network, Linear Regression, Logistic Regression
- Złożoność algorytmów ukryta za stosunkowo prostym API
 - Logika DM dostępna z poziomu kwerend SQL, wyniki analiz – dataset/rowset
 - Aplikacje mogą albo korzystać z SQL Server Data Mining by wykorzystywać wyuczone reguły, albo pokazywać wzorce użytkownikowi
- Jak rozszerzać:
 - „SQL Server Data Mining: Plug-In Algorithms”
 - W skrócie: trzeba zaimplementować 5 interfejsów
 - C++
 - Managed Plug-in Algorithm Framework (April 2006 Feature Pack)
 - C#
- SQL 2008
 - Dane w treningowej i testowej partycji
 - Walidacja krzyżowa dla różnych podzbiorów (automat!)
 - Zapytanie o dodatkowe dane (nie użyte w model); może warto coś jeszcze uwzględnić?
 - Filtrowanie danych przy budowie modelu (a może kobiety i mężczyźni jednak powinni być analizowani oddzielnie?)
 - Łączne wykorzystywanie danych dyskretnych i ciągłych; nawet dla jednej kolumny
 - Zmiany w TimeSeries
 - SQL Server 2005 : ARTXP (MS Research); dobry do krótkich okresów
 - SQL Server 2008: ARIMA (powszechniejszy; lepszy dla długich okresów)

Data Mining i SQL

- Jak połączyć świat SQL i Data Mining

```
CREATE MINING MODEL CreditRisk
(CustID      LONG KEY,
Gender      TEXT DISCRETE,
Income      LONG CONTINUOUS,
Profession  TEXT DISCRETE,
Risk        TEXT DISCRETE
PREDICT)
USING Microsoft_Ddecision_Trees
```

```
INSERT INTO CreditRisk
(CustId, Gender, Income,
Profession, Risk)
Select
CustomerID, Gender, Income,
Profession,Risk
From Customers
```

```
Select NewCustomers.CustomerID, CreditRisk.Risk,
PredictProbability(CreditRisk)
FROM CreditRisk PREDICTION JOIN NewCustomers
ON CreditRisk.Gender=NewCustomer.Gender
AND CreditRisk.Income=NewCustomer.Income
AND CreditRisk.Profession=NewCustomer.Profession
```

Dostęp do danych

Mechanizmy dostępu do danych

- ADO.NET
 - DataReader
 - DataAdapter + DataSet
- LINQ
- Entity Data Model
- Microsoft Synchronization Services
- Service Broker
- Astoria
- Jasper

Microsoft Synchronization Services

- Początki
 - Aplikacja: DataSet + wersje wierszy
 - Serwer: Replikacja (dalej istnieje! I jest rozwijana...)
- SQL Server Change Tracking
 - W środku pojemnika DB (bez ręcznego) timestamp
 - Można: pokaż zmiany od danego punktu w czasie
- SQL Server Conflict Detection
 - Dodawanie, zmiana, oraz usuwanie danych
- ADO.NET + SQL Server CE 3.5 + ADO.NET SyncServices
- Cel: dodanie scenariusza bezpołączeniowego bez konieczności zmiany kodu

Service Broker

- Pierwsza wersja w SQL Server 2005
 - Model asynchroniczny
 - Luźno połączone komponenty bazodanowe
 - Procedura „wołana” w momencie przyjęcia komunikatu
 - Transakcyjne; między serwerami itp
 - De facto – MSMQ oparte o SQL
- SQL Server 2008
 - Priorytety dla komunikatów
 - Oraz – łatwiejsze programowanie!

Projekt LINQ

C# 3.0

VB 9.0

Others

.NET Language Integrated Query

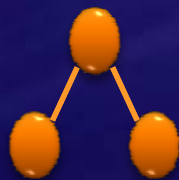
LINQ to
Objects

LINQ to
DataSets

LINQ to
SQL

LINQ to
Entities

LINQ to
XML



Obiekty



Dane
relacyjne

```
<book>  
  <title/>  
  
  <author/>  
  <year/>  
  <price/>  
</book>
```

XML

LINQ - składnia

- Koncepcja wyrażenia lambda (lambda expression)
 - Użycie w wyrażeniach-kwerend (Query Expression)
 - LINQ, DLINQ, XLINQ, (E)LINQ,...
 - Realizacja używając metod-rozszerzeń (Extension Methods)
- Inferencja typów (type inference)
 - Inne zastosowanie niż object
- Inicjalizatory obiektów i kolekcji
- Typy anonimowe (rozbudowa metod anonimowych)

```
gridCategories.DataSource = from category in
db.Production.ProductSubcategory
orderby category.ProductSubcategoryID
select new {
    CategoryID = category.ProductSubcategoryID
    Name = category.ProductSubcategoryName
}
```

```
int[] numbers = { 5, 4, 1, 3, 9, 8, 6, 7, 2, 0 };
int oddNumbers = numbers.Count(n => n % 2 == 1);
```

```
gridProducts.DataSource =
from prod in db.GetAllProductsByCategoryID(1)
orderby prod.ProductID
select prod;
```

Entity Data Model

- W zasadzie O/R mapper, ale
 - Graficzny projektant (relacje itp.)
- Praca z obiektami
- Można stosować specjalne adaptery
 - EDM to SQL Server
 - LINQ to EDM
- A co z DataSet?

Astoria

- Generowanie warstwy DAL jako Web Service
 - Kwerenda za pośrednictwem URI
 - `http://myserver/data.svc/Customers[ALFKI]`
 - `/Orders[Active eq true]?$orderby=OrderDate`
 - Wynik zwracany w XML (lub JSON)
- Tradycyjne programowanie w .NET, ale też Silverlight
- Wykorzystuje Entity Framework
- Więcej informacji:
 - <http://astoria.mslivelabs.com>

Jasper

- Dynamiczne generowanie klas dla danych
 - De facto – O/R mapper w locie
 - Bez mapowania itp
 - Konfigurowalne
- Wykorzystuje Entity Framework
- Automatyczny binding
 - Budowane: ASP.NET, WinForms, WPF
- Więcej informacji:
 - <http://blogs.msdn.com/adonet>

Podsumowanie

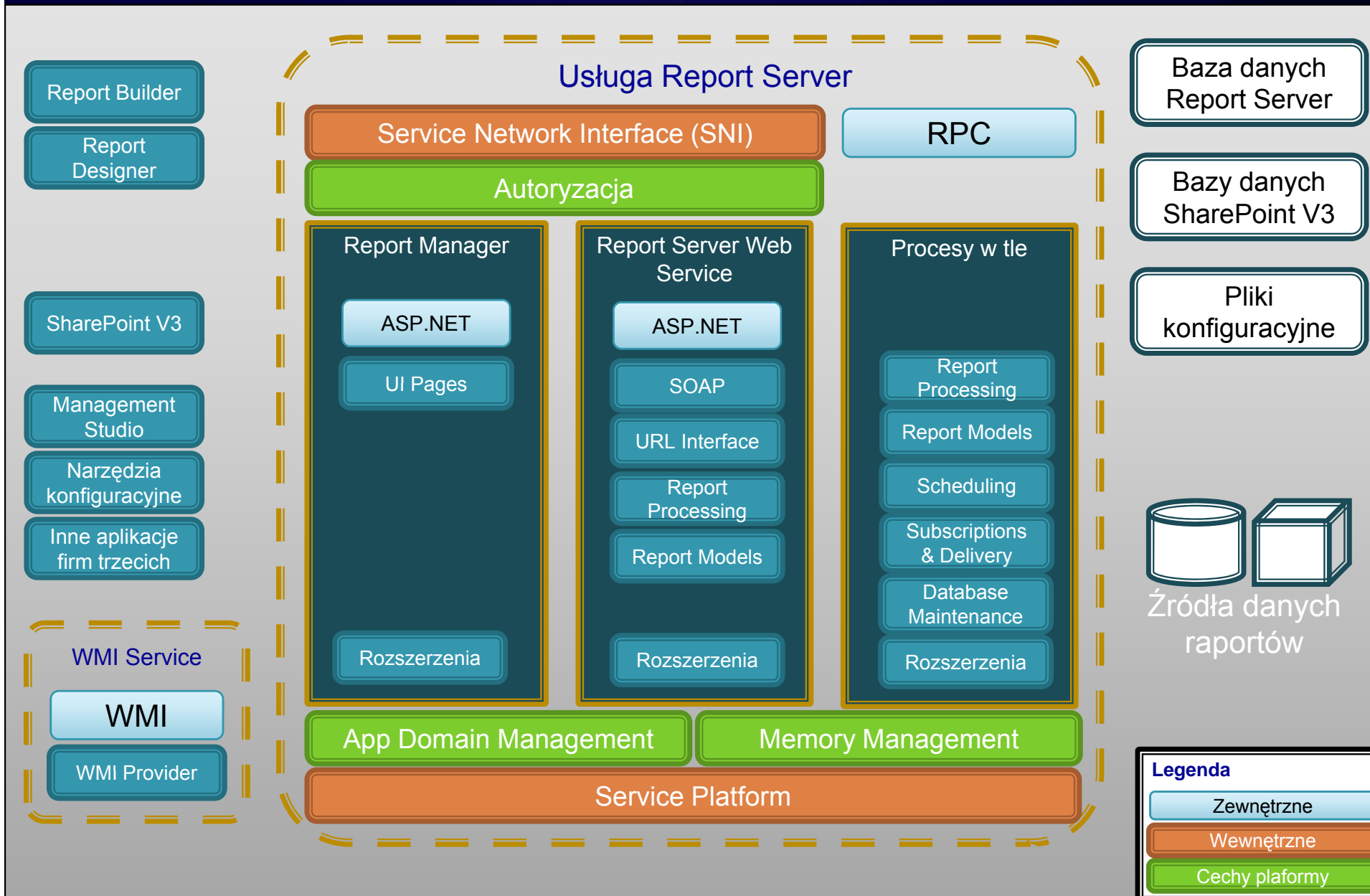
- SQL 2008
 - Motor relacyjny
 - XML
 - FullText
 - Hierarchie
 - Dane binarne
 - OLAP
 - DataMining
 - Narzędzia: wsparcie dla projektów zespołowych
- Dostęp do danych
 - Co najmniej kilka opcji

Pytania?

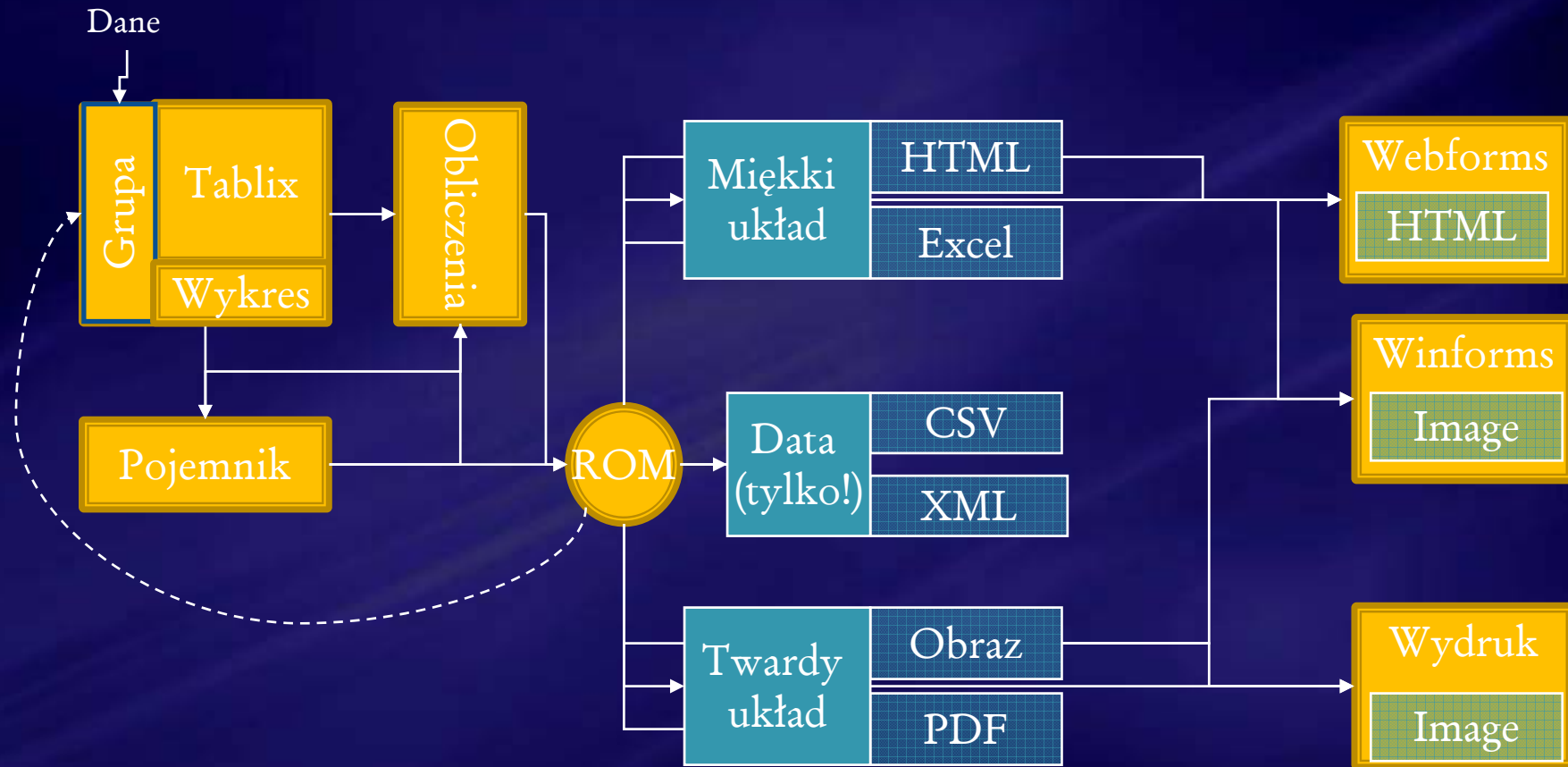
Kontakt: Tomasz Kopacz
tkopacz@microsoft.com

Dalsze dodatkowe slajdy

Architektura Reporting Services 2008



Architektura 2008



Tablix

- Matrix +:
 - Wiele równoległych wierszy/kolumn na dowolnym poziomie
 - Element może być dynamiczny / statyczny
 - Nagłówek opcjonalny
- Tabela +:
 - Dynamiczne, zagnieżdżone grupy
 - Wiele równoległych grup w wierszach
 - Statyczne wiersze
 - Opcjonalnie - nagłówki

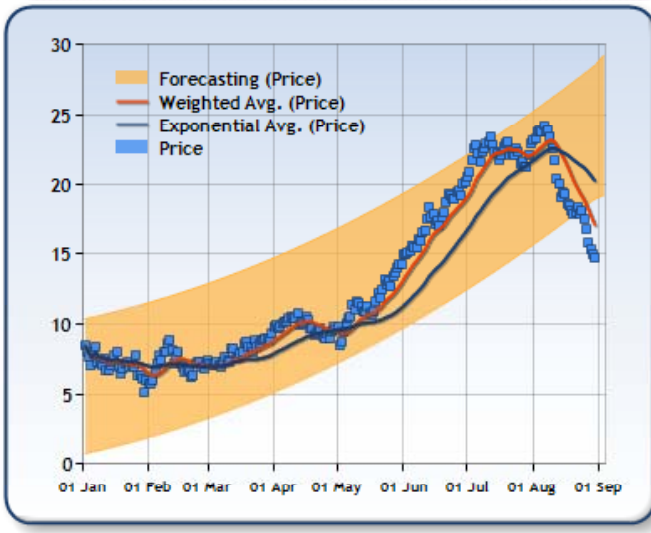
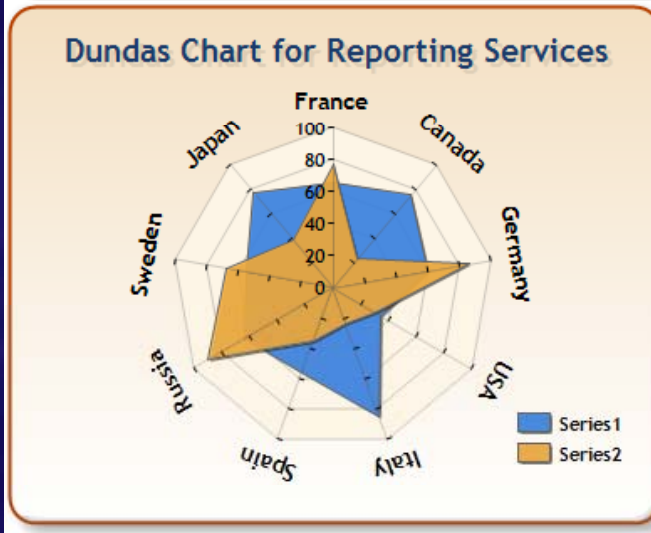
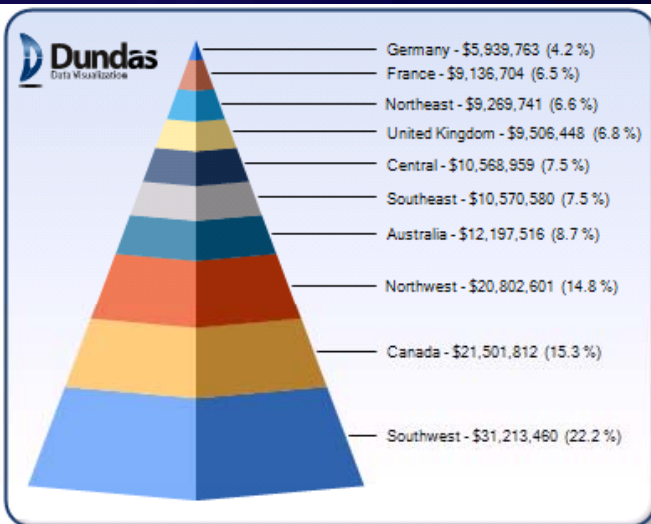
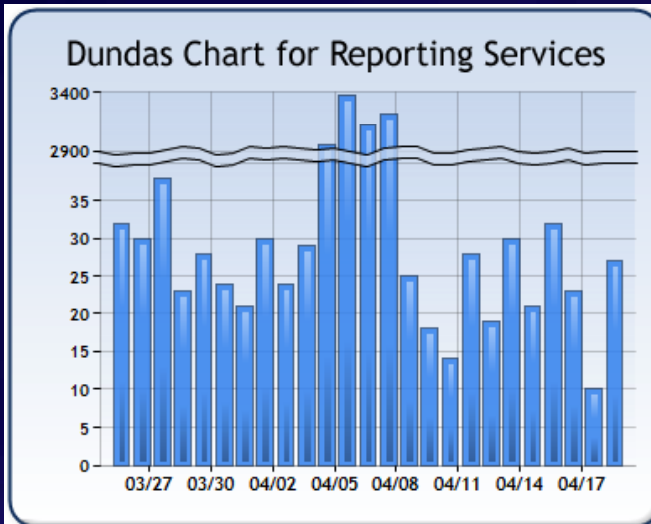
Przykład Tablix

(beznagłówkowe)
Elementy
dodane
dynamicznie

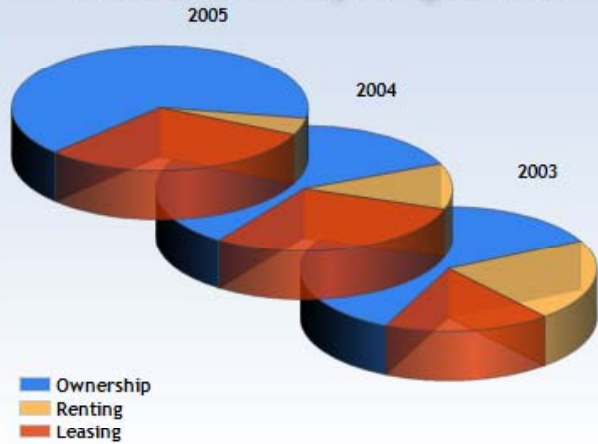
Elementy
statyczne

	Year			Region		Total
	2001	2002	Growth	West	East	
Retail						
Acme	1,115	1,331	19%	1,111	1,225	2,446
Nadir, Inc.	152	642	322%	222	572	794
Wholesale						
ABC Corp.	11,156	13,312	19%	11,111	12,257	24,468
XYZ, Ltd.	1,523	6,421	322%	2,222	5,722	7,944
Grand Total	13,946	21,706	56%	14,666	19,786	35,653

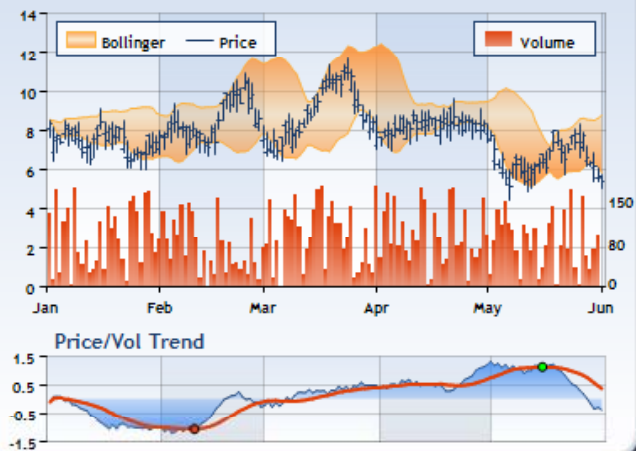
Reporting Services - Dundas



Dundas Chart for Reporting Services



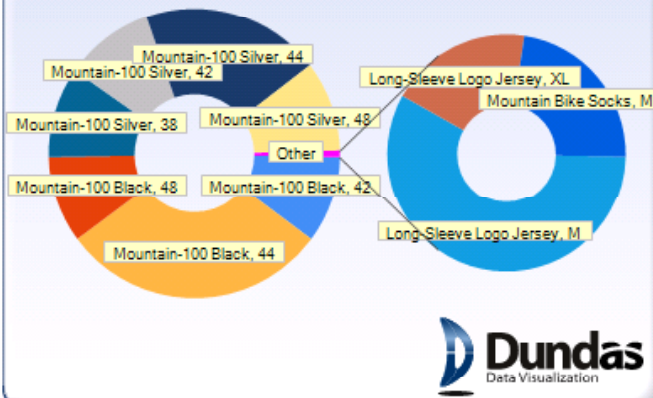
Dundas Chart for Reporting Services



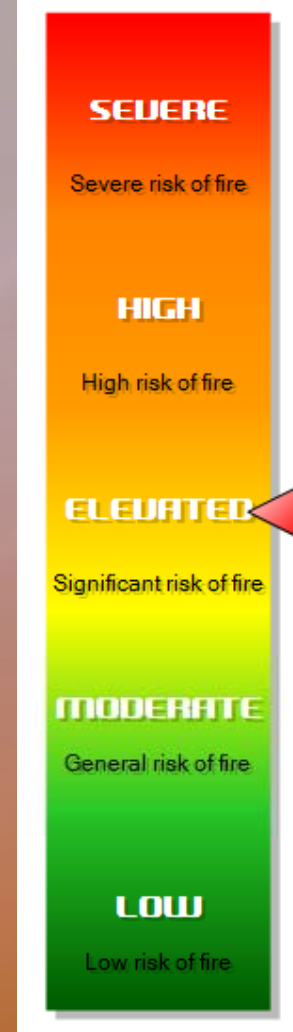
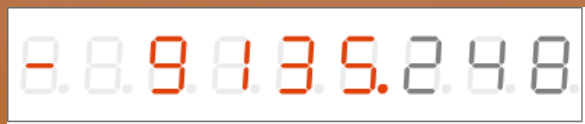
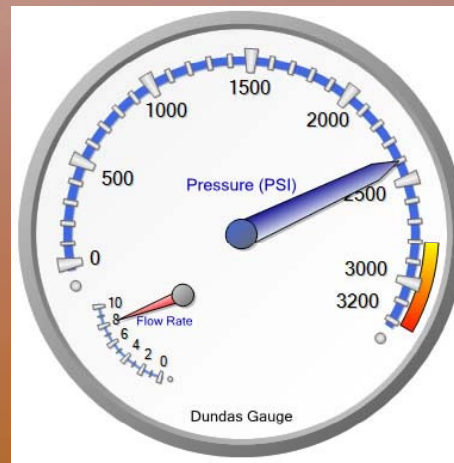
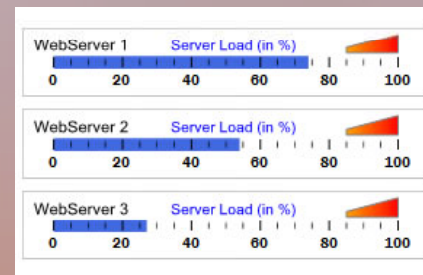
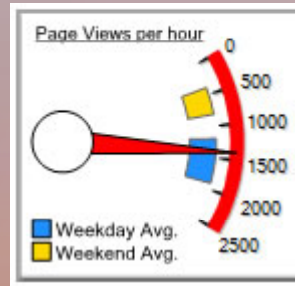
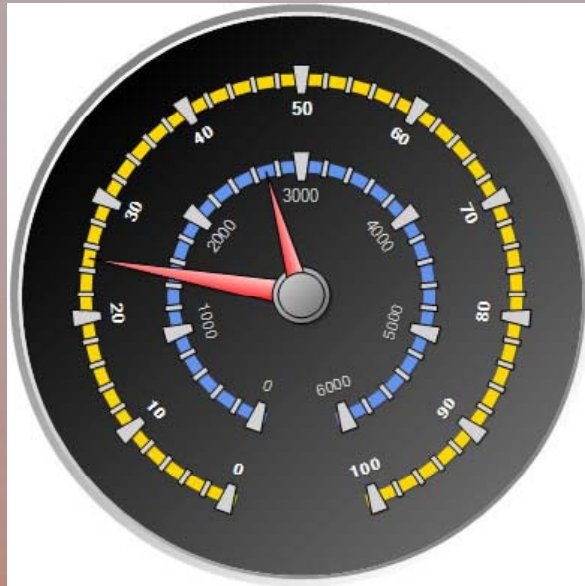
Dundas Chart for Reporting Services



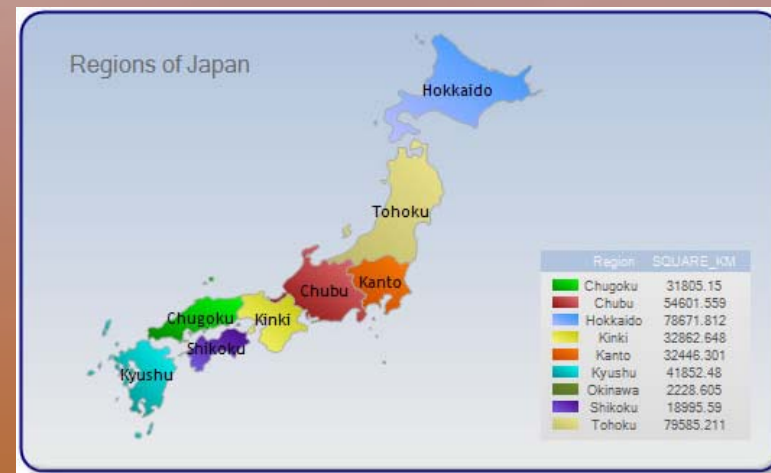
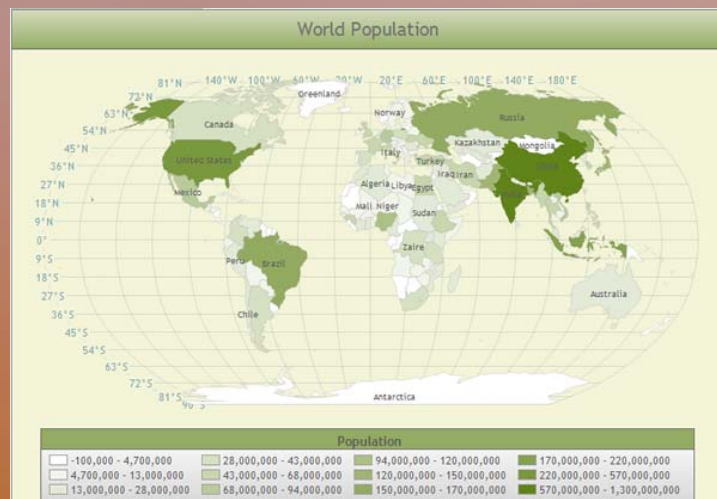
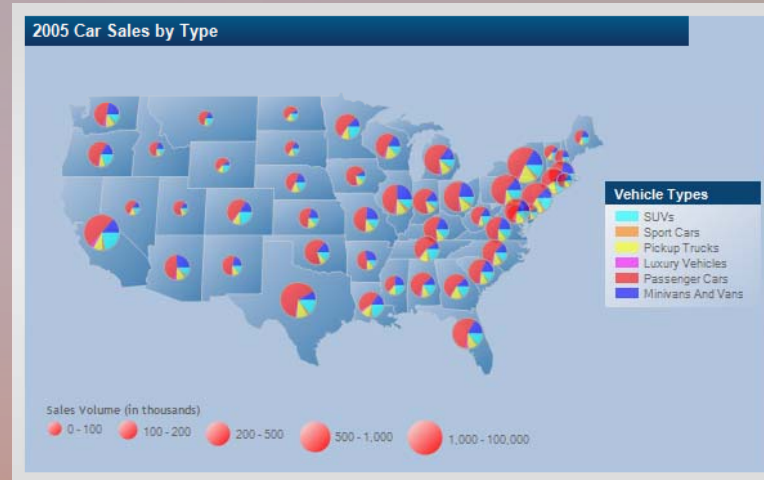
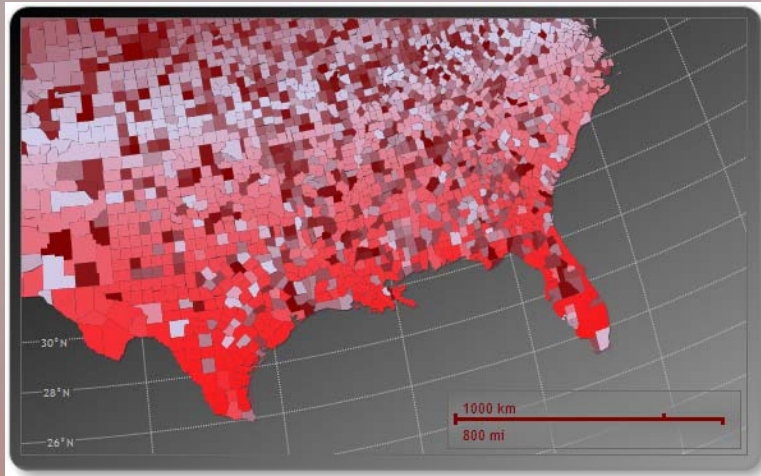
Dundas Chart for Reporting Services



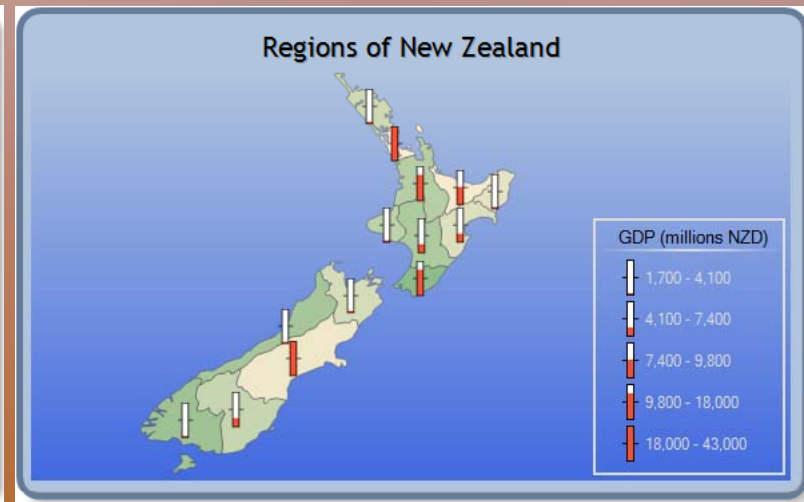
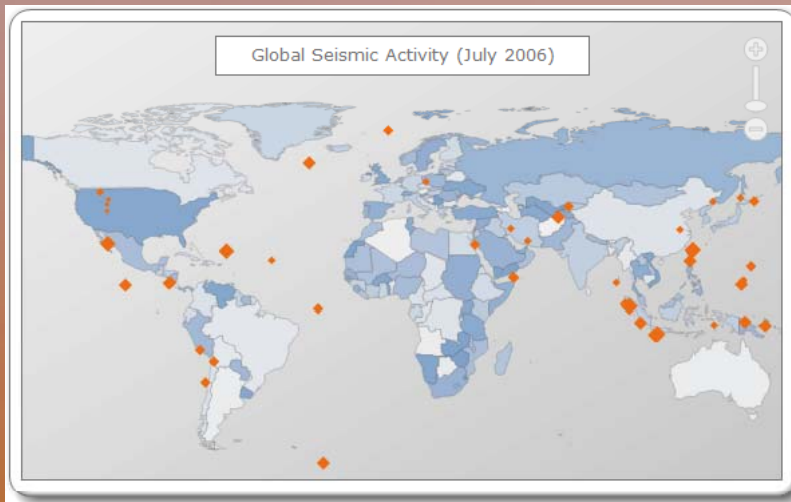
Gauge Examples



Map Examples



Map Examples



Microsoft[®]

Your potential. Our passion.[™]

© 2007 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.