

Marco Russo
Alberto Ferrari

Microsoft SQL Server 2016
Analysis Services

Modelowanie tabelaryczne

Przekład: Jakub Niedźwiedź, Witold Sikorski, Marek Włodarz

APN Promise, Warszawa 2017

Spis treści

<i>Przedmowa</i>	xiii
<i>Wprowadzenie</i>	xv
1 Wprowadzenie do modelu tabelarycznego	1
Modele semantyczne w Analysis Services	1
Czym jest usługa Analysis Services i dlaczego należy jej używać?	1
Krótka historia Analysis Services	3
Model tabelaryczny i wielowymiarowy	4
Model tabelaryczny	4
Model wielowymiarowy	7
Po co są dwa modele?	8
Przyszłość Analysis Services	10
Azure Analysis Services	10
Wybór odpowiedniego modelu dla naszego projektu	11
Licencjonowanie	11
Aktualizacja poprzednich wersji Analysis Services	12
Prostota korzystania	12
Kompatybilność z Power Pivot	13
Kompatybilność z Power BI	13
Cechy wydajności zapytań	13
Cechy wydajności przetwarzania	14
Uwarunkowania sprzętowe	14
BI czasu rzeczywistego	15
Narzędzia klienckie	16
Porównanie funkcji	16
Porównanie DAX i MDX	18
Język DAX	18
Język MDX	19
Wybieranie języka zapytań dla modelu tabelarycznego	20
Wprowadzenie do silników kalkulacyjnych modelu tabelarycznego	20
Wprowadzenie do VertiPaq	21
Wprowadzenie DirectQuery	22
Poziomy zgodności modelu tabelarycznego (1200 kontra 110x)	24

Analysis Services i Power BI	26
Podsumowanie	27
2 Rozpoczynanie pracy w modelu tabelarycznym.....	29
Budowanie środowiska deweloperskiego	29
Komponenty środowiska deweloperskiego	29
Licencje	32
Proces instalacji.....	32
Korzystanie z SQL Server Data Tools	41
Tworzenie nowego projektu.....	41
Konfiguracja nowego projektu.....	43
Import z PowerPivot	49
Importowanie z Power BI	50
Import wdrożonego projektu z Analysis Services.....	51
Zawartość projektu tabelarycznego	51
Budowa prostego modelu tabelarycznego.....	53
Ładowanie danych do tabel.....	53
Praca w widoku diagramu	64
Nawigowanie przy użyciu Tabular Model Explorer	68
Wdrażanie modelu tabelarycznego	70
Odpytywanie modelu tabelarycznego przy użyciu programu Excel	71
Połączenie z modelem tabelarycznym.....	72
Korzystanie z tabel przestawnych.....	75
Korzystanie z fragmentatorów.....	76
Sortowanie i filtrowanie wierszy i kolumn.....	79
Używanie formuł kostek Excela	81
Odpytywanie modelu tabelarycznego przy użyciu Power BI Desktop ...	83
Tworzenie połączenia z modelem tabelarycznym	83
Budowanie podstawowego raportu Power BI.....	85
Dodawanie wykresów i fragmentatorów	86
Interakcja z raportem	88
Korzystanie z SQL Server Management Studio	89
Importowanie z Power Pivot	92
Importowanie z Power BI Desktop	93
Używanie DAX Studio jako alternatywy dla SSMS	93
Podsumowanie	94

3 Ładowanie danych do modelu tabelarycznego	95
Istota źródła danych	95
Personifikacja	97
Poświadczenia po stronie serwera i po stronie klienta	99
Praca z dużymi tabelami	100
Ładowanie z serwera SQL Server	101
Ładowanie z listy tabel	104
Ładowanie z zapytania SQL	107
Ładowanie z widoków	108
Otwieranie istniejących połączeń	109
Ładowanie z programu Access	110
Ładowanie z Analysis Services	111
Korzystanie z edytora MDX	112
Ładowanie z tabelarycznej bazy danych	113
Ładowanie z pliku Excela	116
Ładowanie z pliku tekstowego	118
Ładowanie ze schowka	120
Ładowanie z raportu usług Reporting Services	123
Ładowanie danych przy użyciu źródła danych raportu	123
Ładowanie z raportów przy użyciu strumieniowych źródeł danych ..	129
Ładowanie ze strumieniowych źródeł danych	131
Ładowanie z SharePoint	133
Wybór właściwej metody ładowania danych	133
Podsumowanie	135
4 Wprowadzenie do DAX	137
Wprowadzenie do języka DAX	137
Składnia DAX	139
Typy danych języka DAX	140
Operatory DAX	143
Odwoływanie się do kolumn i miar	145
Funkcje agregujące	145
Funkcje tablicowe	146
Kontekst wykonania	147
CALCULATE i CALCULATETABLE	149
Zmienne	152
Miary	152
Kolumny obliczane	154

Tabele obliczane	156
Tworzenie zapytań w języku DAX	158
Formatowanie kodu DAX	159
DAX Formatter, DAX Studio i DAX Editor	161
Podsumowanie	162
5 Tworzenie hierarchii	163
Hierarchie podstawowe	163
Czym są hierarchie?	163
Kiedy budować hierarchie	165
Tworzenie hierarchii	166
Najlepsze praktyki projektowania hierarchii	167
Hierarchie obejmujące wiele tabel	168
Naturalne i nienaturalne hierarchie	170
Hierarchie rodzic-dziecko	171
Czym są hierarchie rodzic-dziecko?	171
Konfigurowanie hierarchii rodzic-dziecko	172
Operatory jednoargumentowe	177
Podsumowanie	182
6 Modelowanie danych w modelu tabelarycznym	183
Różne techniki modelowania danych	184
Korzystanie z bazy danych OLTP	186
Praca z modelami wymiarowymi	187
Praca z powoli zmieniającymi się wymiarami	188
Praca ze zdegenerowanymi wymiarami	192
Korzystanie z migawkowych tabel faktów	193
Korzystanie z widoków w celu odseparowania się od bazy danych	195
Typy relacji	197
Kardynalność relacji	198
Propagowanie filtru w relacjach	204
Stan aktywny relacji	208
Implementowanie relacji w języku DAX	210
Normalizacja kontra denormalizacja	211
Kolumny obliczane kontra zewnętrzny proces ETL	215
Odwołania cykliczne przy korzystaniu z tabel obliczanych	218
Podsumowanie	219

7	Tabular Model Scripting Language (TMSL)	221
	Definiowanie obiektów w TMSL.....	221
	Obiekt Model	223
	Obiekt DataSource.....	226
	Obiekt Table	228
	Obiekt Relationship	235
	Obiekt Perspective	237
	Obiekt Culture	238
	Obiekt Role	240
	Polecenia TMSL	242
	Operacje na obiektach w TMSL.....	243
	Operacje odświeżania danych i zarządzania bazą danych w TMSL ..	244
	Skrypty TMSL	245
	Podsumowanie	246
8	Warstwa prezentacji modelu tabelarycznego	247
	Ustawianie metadanych dla tabeli kalendarzowej.....	247
	Nazwy, sortowanie i formatowanie	248
	Nazywanie obiektów	248
	Ukrywanie kolumn i miar	251
	Porządkowanie miar i kolumn	251
	Sortowanie danych w kolumnie	253
	Formatowanie.....	256
	Perspektywy	259
	Właściwości związane z Power View	262
	Domyślny zestaw pól.....	262
	Właściwości sterujące zachowaniem tabeli	264
	Kluczowe wskaźniki wydajności (KPI).....	265
	Translacje	268
	Tworzenie pliku translacji	270
	Zapisywanie tłumaczeń nazw w pliku translacji	271
	Wybieranie edytora dla plików translacji.....	273
	Importowanie pliku translacji.....	275
	Testowanie tłumaczeń przy użyciu narzędzi klienckich.....	276
	Usuwanie tłumaczenia.....	277
	Najlepsze praktyki dotyczące tłumaczeń.....	278
	Wybieranie ustawień językowych w modelu tabelarycznym.....	279

Zmianianie ustawień językowych przy korzystaniu ze zintegrowanego obszaru roboczego	281
Zmianianie ustawień językowych przy korzystaniu z serwera obszaru roboczego	282
Podsumowanie	282
9 Korzystanie z DirectQuery	283
Konfigurowanie DirectQuery	284
Ustawianie DirectQuery w środowisku deweloperskim	284
Włączanie trybu DirectQuery po wdrożeniu	291
Ograniczenia modelu tabelarycznego dla DirectQuery	294
Obsługiwane źródła danych	294
Ograniczenia dla źródeł danych	294
Ograniczenia modelowania danych	295
Ograniczenia dotyczące formuł DAX	295
Ograniczenia występujące w formułach MDX	297
Dostrajanie ograniczeń zapytań	298
Wybieranie pomiędzy trybem DirectQuery a VertiPaq	300
Podsumowanie	301
10 Bezpieczeństwo	303
Uwierzytelnianie użytkowników	303
Łączenie się z serwerem Analysis Services spoza domeny	304
Kerberos i problem podwójnego przeskoku	305
Role	306
Tworzenie ról bazodanowych	307
Członkostwo w wielu rolach	309
Zabezpieczenia administracyjne	310
Przyznawanie uprawnień poprzez rolę Server Administrator	310
Role bazodanowe i uprawnienia administracyjne	311
Zabezpieczenia danych	312
Podstawowe zabezpieczenia danych	313
Testowanie ról zabezpieczeń	314
Zaawansowane wyrażenia filtra wierszy	319
Zabezpieczenia w kolumnach i tabelach obliczanych	324
Korzystanie z tabeli uprawnień	325
Szacowanie wpływu zabezpieczeń danych na wydajność	326
Tworzenie zabezpieczeń dynamicznych	327

Funkcje DAX związane z zabezpieczeniami dynamicznymi	327
Implementowanie zabezpieczeń dynamicznych przy użyciu CUSTOMDATA	328
Implementowanie zabezpieczeń dynamicznych przy użyciu USERNAME	329
Zabezpieczenia w trybie DirectQuery	332
Zabezpieczenia i personifikacja w DirectQuery	332
Zabezpieczenia na poziomie wierszy w wersjach SQL Server wcześniejszych niż 2016	334
Monitorowanie zabezpieczeń	335
Podsumowanie	337
11 Przetwarzanie i partycjonowanie modelu tabelarycznego	339
Automatyzowanie wdrożenia na serwerze produkcyjnym.	339
Partycjonowanie tabel	341
Definiowanie strategii partycjonowania	341
Definiowanie partycji tabeli w modelu tabelarycznym	344
Zarządzanie partycjami tabeli	348
Opcje przetwarzania	351
Dostępne operacje przetwarzania	352
Definiowanie strategii przetwarzania	358
Wykonywanie przetwarzania	361
Automatyzowanie przetwarzania	365
Korzystanie z poleceń TMSL	366
Korzystanie z SQL Server Integration Services	374
Korzystanie z bibliotek Analysis Management Objects (AMO) i Tabular Object Model (TOM)	377
Korzystanie z PowerShell	380
Przykładowe skrypty przetwarzania	381
Przetwarzanie bazy danych	381
Przetwarzanie tabel	382
Przetwarzanie partycji	383
Partycje kroczące	384
Podsumowanie	388
12 Wewnątrz VertiPaq	389
Struktury VertiPaq	390
Istota magazynu kolumnowego	390
Kodowanie wartości kontra kodowanie skrótów	394

Kompresja RLE	397
Kontrolowanie kodowania kolumn	399
Hierarchie i relacje	400
Segmentacja i partycjonowanie	402
Odczytywanie wewnętrznych metadanych VertiPaq	404
Korzystanie z DMV w celu sprawdzenia użycia pamięci przez VertiPaq	404
Interpretowanie raportów VertiPaq Analyzer	406
Użycie pamięci VertiPaq	410
Użycie pamięci na dane	410
Użycie pamięci podczas przetwarzania	412
Użycie pamięci w zapytaniach	413
Opcje przetwarzania	414
Co dzieje się podczas przetwarzania	415
Dostępne opcje przetwarzania	416
Podsumowanie	418
13 Programowy interfejs modelu tabelarycznego	419
Wprowadzenie do bibliotek AMO oraz TOM	419
Wprowadzenie do AMO	420
Wprowadzenie do TOM	422
Wprowadzenie do poleceń TMSL	428
Programowe tworzenie bazy danych	430
Automatyzowanie odświeżania danych i partycjonowania	434
Analizowanie metadanych	434
Manipulowanie modelem danych	437
Automatyzowanie wdrożenia projektu	439
Kopiowanie tej samej bazy danych na różne serwery	439
Wdrażanie pliku model.bim z wyborem nazwy bazy danych i serwera	440
Podsumowanie	442
14 Monitorowanie i dostrajanie usługi tabelarycznej	443
Znajdowanie procesu usług Analysis Services	443
Zasoby zużywane przez Analysis Services	445
CPU	445
Pamięć	447
Operacje I/O	449

Zrozumienie konfiguracji pamięci	450
Korzystanie z liczników wydajnościowych związanych z pamięcią	455
Korzystanie z dynamicznych widoków zarządzania	460
Widoki DMV przydatne do monitorowania usługi tabelarycznej	462
Automatyzowanie gromadzenia informacji monitorowania i dzienników	464
Liczniki wydajności	464
SQL Server Profiler	465
ASTrace	469
Flight Recorder	469
Zdarzenia rozszerzone	470
Inne narzędzia komercyjne	471
Monitorowanie odświeżania danych (przetwarzania)	471
Monitorowanie zapytań	475
Podsumowanie	478
15 Optymalizowanie modeli tabelarycznych	479
Optymalizowanie użycia pamięci	479
Usuwanie nieużywanych kolumn	480
Zmniejszanie wielkości słownika	480
Wybieranie typu danych	484
Redukowanie wielkości bazy danych poprzez wybór kolejności sortowania	486
Ulepszanie kodowania i rozmiaru bitowego	489
Optymalizowanie wielkich wymiarów	490
Projektowanie modeli tabelarycznych dla wielkich baz danych	495
Optymalizowanie kompresji poprzez podział kolumny	495
Optymalizowanie czasu przetwarzania wielkich tabel	497
Agregowanie tabel faktów na różnych poziomach szczegółowości ..	498
Projektowanie modeli tabelarycznych dla rozwiązań niemal czasu rzeczywistego	503
Wybieranie pomiędzy DirectQuery a VertiPaq	503
Wykorzystanie partycji	504
Redukowanie czasu ponownych obliczeń	507
Zarządzanie blokadami podczas przetwarzania	509
Podsumowanie	510

16 Wybieranie sprzętu i wirtualizacja	511
Wymiarowanie sprzętu	511
Zegar i model procesora	513
Prędkość i wielkość pamięci	516
Architektura NUMA	517
Dysk i I/O	519
Wymagania sprzętowe dla DirectQuery	520
Optymalizowanie konfiguracji sprzętowej	521
Ustawienia zasilania	521
Hyper-threading	523
Ustawienia NUMA	523
Wirtualizacja	524
Dzielenie węzłów NUMA pomiędzy różne maszyny wirtualne	524
Przydzielanie pamięci dla maszyny wirtualnej	525
Skalowalność rozwiązania tabelarycznego SSAS	526
Skalowalność dla pojedynczej bazy danych (wielki rozmiar)	526
Skalowalność dla wielkiej liczby użytkowników	527
Podsumowanie	527
<i>Indeks</i>	529
<i>O autorach</i>	556