

Obsah

Úvod	9
Kapitola 1	
Business Intelligence, datové sklady	11
Přechod od transakčních databází k analytickým	13
Kvalita údajů pro analýzy	14
Nevhodnost transakčních databází pro analýzy	14
Zpracování dat z operačního prostředí	16
Datový sklad	17
Datové trhy	21
Metody budování datového skladu	21
Kapitola 2	
Architektura, modelování a implementace procesů	
Business Intelligence na MS SQL Serveru 2005	27
Architektura analytických služeb MS SQL Serveru 2000	28
Architektura analytických služeb MS SQL Serveru 2005	28
XML for Analysis (XMLA)	29
Metoda Discover	29
Metoda Execute	31
Podpora XMLA na platformě analytických služeb MS SQL Serveru 2005	32
Výhody a nevýhody	33
Modelování procesů Business Intelligence	33
Časová posloupnost procesu Business Intelligence	37
Business Intelligence v reálném čase	38
Nástroje pro práci s analytickými službami	39
Implementace a přístup k UDM přes Business Intelligence Development Studio	39
Data Source Views	41
SQL Server Management Studio	42
Dotazy v prostředí SQL Server Management Studio	45
Výpis parametrů a vlastností	47
Využití šablon skriptů	48
Kapitola 3	
Integrační služby v MS SQL Serveru 2005	51
Extrakce, transformace a přenos	54

Oblast vynášení dat	55
Extrakce	56
Čištění dat	56
Transformace	57
Problémy při transformaci dat	58
Přenos	60
Chyby a problémy procesu ETL	60
Testování etapy ETL	60
Architektura integračních služeb MS SQL Serveru 2005	61
Vytvoření transformačního balíčku pomocí průvodce	63
Námět příkladu	64
Vytvoření projektu	64
Seznámení s návrhovým prostředím pro návrh integračních služeb	71
Záložka ControlFlow	72
Kontejnery integračních služeb	73
For Loop Container	73
Foreach Loop Container	74
Sequence Container	75
Úlohy pro přípravu dat	75
Workflow úlohy	76
Úlohy pro SQL Server	76
Skriptové úlohy	77
Úlohy využívající analytické služby	77
Execute SQL Task	79
Data Flow Task	83
Aggregate	85
Audit	87
Conditional Split	88
Copy Column	91
Data Conversion	92
Derived Column	92
OLE DB Source Editor	96
OLE DB Destination Editor	98
Data Flow Path Editor	99
Import dat z textového souboru	103
Alternativní řešení	116
Pokročilejší konverze dat	117
Analýza problému	117
Návrh integračního projektu	118
Operátory pro konstruování SSIS podmínek	119
Funkce pro konstruování SSIS podmínek	119
Rozdělení dat na platná a neplatná	121
Výběr dat ze dvou datových zdrojů, které obsahují duplicitní data	125
Definování připojení na zdroje dat	127

Definice připojení pro ukládání výstupních hodnot a duplicitních záznamů	128
Definice připojení pro ukládání informací o chybách	128
Definice připojení k databázi AdventureWorksDW	128
Control Flow	130
Data Flow	133
Odstranění exaktních duplicit	134
Odstranění „neostrých“ duplicit	136
Uložení odhalených duplicit	140
Ladění projektu	142
Nasazení integračních projektů	145

Kapitola 4

Reportovací služby	149
Enterprise Reporting	150
Architektura a filozofie reportovacích služeb	150
Report Manager	152
Životní cyklus reportu	152
Návrh reportu	153
Správa reportu	155
Doručení reportu	155
Architektura reportovacích služeb	155
Struktura jazyka RDL pro návrh reportů	156
Grafické prvky pro návrh reportů	157
Textbox	158
Line	159
Rectangle	160
Image	160
Prvky pro zobrazování dat	163
Výrazy v jazyce RDL	166
Agregační funkce	167
Návrh výstupní sestavy dat z databáze	169
Návrh dotazu pro výběr dat	171
Výběr zdroje dat	174
Návrh reportu pomocí průvodce	177
Spuštění reportu ve vývojovém prostředí	182
Návrh reportu ve vývojovém prostředí	183
Grafické prostředí pro návrh reportu	186
Pohled „za oponu“ reportovacích služeb	190
Nástroj pro konfiguraci reportovacích služeb	191
Ovládací prvky zobrazení reportu	192
Možnosti exportu výstupní sestavy	194
Návrh složitějšího reportu	198
Definice množiny dat pro report	198
Návrh reportu	200

Grafické úpravy reportu v prostředí MS Visual Studio	205
Interaktivní parametrický report	207
Report Builder	212
Návrh reportu pomocí Report Builderu	218
Správa reportu	222
Report Manager	222
Doručení reportu	227
Data – Driven Subscription	229

Kapitola 5

OLAP analýzy	233
Zopakování základů teorie relačních databází	233
Vztahy mezi entitami	236
Normalizace databází	237
Multidimenzionální databáze	239
Multidimenzionální databázový model	240
Porovnání relačního a multidimenzionálního modelu	241
Teoretický úvod do problematiky OLAP	242
Fakty a dimenze	243
Schémata tabulek dimenzí	245
Úložiště multidimenzionálních dat MOLAP, ROLAP, HOLAP	246
Úvodní příklad pro OLAP analýzu	248
Databázové tabulky pro cvičný příklad	254
Relační databáze pro cvičný příklad	254
Multidimenzionální databáze pro cvičný příklad	256
Vytvoření OLAP kostky rodinného rozpočtu pomocí průvodce	259
Definice datových zdrojů	260
Definice pohledů na datové zdroje	263
Vytvoření kostky	267
OLAP kostka v jazyce SQL – klauzule CUBE	276
Vytvoření OLAP kostky z datového skladu	289
Práce s OLAP kostkou v prostředí BI Dev Studio	297
Cube Structure	297
Dimension Usage	298
Perspektivy	300
Particie	301
Browser	304
Kalkulace (záložka Calculations)	305
Klíčové indikátory (záložka KPIs)	306
Záložka Translations	309

Kapitola 6

Dolování dat (data mining)	313
Co data mining neumožňuje	314
Teoretický úvod – stručné statistické minimum	315
Rozdělení pravděpodobnosti a testování hypotéz	315
Statistické metody využívané data miningovými modely	317
Model procesu data miningu	318
Procesní schéma data miningu	319
Učící fáze	320
Analýza a predikce nových případů	320
Algoritmy pro data mining	321
Vicerozměrné shlukové diagramy	322
Nevyvážené rozhodovací stromy	324
Typické okruhy úloh a výběr algoritmů pro jejich řešení	327
Úvodní příklad pro data mining	329
Návrh data miningového modelu	334
Predikce	344
Příklad z praxe – rozlišení jedlých a jedovatých hub	348
Příprava vstupních dat	348
Projekt data miningového modelu	350
Příklad predikce na základě výsledků analýzy	354

Kapitola 7

Klientský přístup k analytickým službám	359
Kategorizace klientského přístupu	359
Tlustý klient	360
Tenký klient	360
Office jako klient analytických služeb	361
Kontingenční tabulka (Pivot Table)	362
Kontingenční tabulka (Pivot Table) v prostředí programu MS Excel	363
Microsoft Excel 2003 jako klient analytických služeb	366
Práce s daty OLAP v režimu off-line	367
Microsoft Excel 2007 jako klient analytických služeb	369

Příloha 1

Cvičná databáze Adventure Works	373
Instalace cvičné databáze Adventure Works pomocí skriptů	381

Příloha 2

Cvičná databáze Adventure Works DW	383
Instalace cvičné databáze Adventure Works DW pomocí skriptů	384