Содержание

Об авторах 14

Введение 15

Для кого предназначена книга 15
О чем эта книга 16
Структура книги 17
Что необходимо для использования данной книги 19
Соглашения, принятые в этой книге 19
Исходный код 19
Ждем ваших отзывов! 20

ЧАСТЬ I. Основы 21

Глава 1. Введение в организацию хранилищ данных и Analysis Services 2005 22

Организация хранилищ данных 23
Ключевые элементы хранилища данных 25
Таблица фактов 26
Таблицы измерений 28
Измерения 29
Кубы 30
Схема “звезда” 31
Схема “снежинка” 32
Инмон против Кимболла: различные подходы 33
Бизнес-аналитика — это анализ данных 34
Analysis Services 2005 35
Унифицированная модель измерений 38
Резюме 41

Глава 2. Первое знакомство с Analysis Services 2005 42

Различия между Analysis Services 2000 и Analysis Services 2005 43
Инструменты разработки и администрирования, а также клиентские инструменты 43
Отличия версий Analysis Services 44
Обновление до версии Analysis Services 2005 45
Использование Business Intelligence Development Studio 53
Создание проекта в Business Intelligence Development Studio 53
Создание базы данных Analysis Services с помощью Business Intelligence Development Studio 57
Использование SQL Server Management Studio 76
Панель Object Explorer 77
Создание запросов с использованием редактора запросов MDX 80
Резюме 83

Глава 3. Введение в MDX 84
Что такое MDX 84
Фундаментальные концепции 85
Члены 86
Ячейки 89
Кортежи 91
Наборы 92
Запрос MDX 93
Инструкция SELECT и определение оси 94
Предложение FROM и определение куба 95
Предложение WHERE и определение среза данных 95
Предложение WITH и вычисляемый член 97
Выражения MDX 101
Операторы 103
Арифметические операторы 103
Операторы набора 103
Операторы сравнения 104
Логические операторы 104
Специальные операторы MDX — фигурные скобки, запятые и двоеточия 105
Функции MDX 105
Категории функций MDX 106
Функции набора 107
Функции члена измерения 110
Числовые функции 111
Функции измерения, функции уровня и функции иерархии 112
Функции обработки строковых значений 112
Другие функции 113
Резюме 113

Глава 4. Работа с источниками данных и представлениями источников данных 114
Источник данных 114
Источники данных, поддерживаемые в Analysis Services 116
Сравнение провайдеров данных OLE DB и .NET 117
Представления источников данных 120
Конструктор представления источника данных 120
Рассмотрим подробно представления источников данных 127
Свойства представления источника данных 129
Различные макеты в представлениях источника данных 131
Проверка представления источника данных и начальный анализ данных 131
Несколько источников данных в одном представлении источника данных 134
Резюме 135

Глава 5. Проектирование измерений 137
Работа с мастером измерений 137
Работа с коструктором измерения 146
Атрибуты 147
Иерархии и уровни 150

Просмотр измерений 151
Сортировка членов уровня 157
Оптимизация атрибутов 159
Установка в измерениях параметров перевода на другие языки 160
Создание измерения в соответствии со схемой “снежинка” 162
Создание временного измерения 163
Создание иерархии “предок-потомок” 167
Резюме 169

Глава 6. Проектирование кубов 170

Унифицированная модель измерений 170
Создание куба с помощью мастера 173
Просмотр кубов 180
Измерения куба 185
Типы измерений 186
Просмотр ссылающихся измерений 190
Размерности и размерные группы 192
Вычисляемые члены 199
Вычисляемые размерности 200
Извлечение вычисляемых размерностей 203
Создание проекций 204
Создание переводов 206
Просмотр проекций и переводов 206
Резюме 209

ЧАСТЬ II. Сложные темы 211

Глава 7. Сложные аспекты работы с MDX 212

Основы вычислений 213
Сценарии MDX 213
Ограничение пространства куба/Срез данных куба 229
Использование инструкции SCOPE 230
Использование инструкций CREATE SUBCUBE и DROP SUBCUBE 230
Использование функции EXISTS 232
Использование функции EXISTING 233
Использование вложенного предложения SELECT 233
Удаление пустых ячеек 235
Фильтрация членов по осям 237
Упорядочение и сортировка 238
Пример 1 238
Пример 2 238
Пример 3 239
Пример 4 239
Пример 5 239
Используйте параметризованные запросы 240
Новые функции MDX 241
Резюме 243

Глава 8. Сложные приемы проектирования измерений 244

Пользовательские методы сворачивания данных 245
Модернизация иерархий предок-потомок 253

Унарные операторы 254
Задание имен уровней иерархии “предок-потомок” 258
Использование свойств для настройки измерений 261
Упорядочение членов измерения 261
Член All, заданный по умолчанию член и член Unknown 261
Настройка свойств обработки ошибок 263
Режим хранения 264
Группирование членов для формирования единого члена 265
Добавление сложной аналитики в измерения с помощью мастера Business Intelligence Wizard 266
Сложная аналитика для счетов 266
Добавление сложной аналитики для временн ого измерения 271
Добавление сложной аналитики для измерения 274
Временное измерение сервера 276
Обратная запись измерений 281
Резюме 284

Глава 9. Сложные аспекты проектирования кубов 286

Размерные группы и размерности 286
Добавление и модификация измерений 291
Измерение с отношением фактов 292
Измерение с отношением “многие ко многим” 294
Измерение с отношением добычи данных 296
Ролевые измерения 297
Добавление вычислений в куб 299
Ключевые показатели эффективности (KPI) 306
Создание ключевого показателя эффективности 307
Сквозная детализация данных 315
Действия 316
Типы действий 316
Действие URL 318
Действия отчетов 323
Действие сквозной детализации данных 325
Добавление сложной аналитики в куб 330
Полуаддитивные размерности 330
Конвертирование валюты 332
Работа с разделами 338
Создание локального раздела 339
Создание удаленного раздела 342
Режимы и настройки хранения 350
Создание агрегаций 350
Установка параметров защиты данных 353
Резюме 357

Глава 10. Расширение сценариев MDX с использованием внешних функций 358

Определяемые пользователем функции COM 359
Встроенные определяемые пользователем функции COM 359
Создание определяемых пользователем функций COM 360
Добавление определяемых пользователем функций COM в базу данных Analysis Services 362

Доступ к определяемым пользователем функциям COM в запросах MDX 363
Определяемые пользователем функции .NET (хранимые процедуры) 365
Создание хранимых процедур 365
Добавление хранимых процедур 370
Запрашивание хранимых процедур 373
Отладка хранимых процедур 374
Определяемые пользователем функции COM и сборки .NET 376
Резюме 377

ЧАСТЬ III. Администрирование и настройка производительности 379

Глава 11. Обновление данных UDM 380

Обновление данных измерения в унифицированной модели измерений 381
Добавление члена в измерение 383
Изменение данных для членов измерения 388
Удаление членов измерения 391
Обновление данных куба 393
Обновление значения отдельной ячейки 398
Обновление значений неконцевых ячеек с использованием предложения ALLOCATION 401
Резюме 406

Глава 12. Администрирование Analysis Services 407

Администрирование с использованием инструментов SQL Server 2005 407
Управление сервером 408
Управление объектами Analysis Services 410
Обработка объектов базы данных Analysis Services 411
Управление разделами 421
Управление сборками 425
Создание и восстановление резервной копии 428
Синхронизация 431
Управление защитой 436
Интерактивный режим 438
Администрирование программным методом 442
Модель Analysis Management Objects (AMO) 442
Резервное копирование и восстановление 449
Добавление сборок в Analysis Services 450
HTTP-подключения к Analysis Services 453
Существующие приложения DSO 454
Analysis Services и кластеризация для обработки отказов 456
Резюме 458

Глава 13. Оптимизация производительности 459

Оптимизация модели UDM 462
Тонкая настройка измерений 462
Тонкая настройка куба 468
Оптимизация обработки 476
Создание разделов для ускорения обработки 479
Выбор подходящего типа данных 479

Установка SQL Server и Analysis Services 480
Оптимизация реляционного источника данных 480
Как избежать слишком большого количества агрегаций 481
Использование инкрементной обработки 481
Параллельная обработка 483
Идентификация "узких мест" 486
Повышение производительности запросов 487
Агрегации 489
Создание агрегаций 491
Построение агрегаций на основе частоты использования 499
Параметры построения агрегаций 505
Оптимизация запросов MDX 510
Оптимизация масштабируемости 515
Настройка свойств конфигурации сервера 515
Расширение системы 517
Наращивание системных ресурсов 517
Обработка больших измерений 518
Использование SQL Profiler для анализа производительности 518
Резюме 523

Глава 14. Технология Data Mining 524

Процесс добычи данных 524
Изучение предметной области 527
Изучение и конфигурирование данных 528
Выбор правильного алгоритма 528
Обучение, анализ и прогнозирование 529
Реальные приложения Data Mining 529
Выявление мошенничества 530
Увеличение прибылей в розничной продаже 530
Анализ данных в NBA 530
Data Mining в центрах телефонного обслуживания 531
Алгоритмы Data Mining в Analysis Services 2005 531
Microsoft Decision Trees 533
Naïve Bayes 533
Microsoft Clustering 533
Sequence Clustering 534
Правила ассоциаций 534
Microsoft Neural Networks 534
Time Series 535
Microsoft Linear Regression 535
Microsoft Logistic Regression 535
Работа с моделями Data Mining 536
Реляционная модель Data Mining 536
Модели добычи данных OLAP 559
Анализ куба с помощью измерения Data Mining 564
Резюме 567

Глава 15. Анализ кубов с помощью клиентских компонентов Office 568

Сводные таблицы Microsoft Excel 569
Создание сводной таблицы на основе данных Analysis Services 569

Анализ данных с помощью сводных таблиц 575
Отображение подробных данных 583
Просмотр нескольких размерностей с помощью сводной таблицы 586
Пользовательское группирование данных в сводной таблице 586
Форматирование отчета сводной таблицы 589
Создание сводных диаграмм 591
Создание локальных кубов Excel 593
Сводные таблицы OWC 597
Анализ данных UDM с помощью Microsoft Data Analyzer 604
Резюме 611

Глава 16. Integration Services 613

Создание проекта Integration Services 614
Задачи Integration Services 614
Преобразования Integration Services 614
Создание пакетов Integration Services для операций в Analysis Services 615
Задача Analysis Services Execute DDL Task 615
Обработка объекта Analysis Services 625
Загрузка данных в раздел Analysis Services 627
Задачи Integration Services для Data Mining 636
Автоматизация пакетов SSIS 636
Резюме 641

ЧАСТЬ IV. Сценарии 643

Глава 17. Reporting Services 644

Конструктор отчетов 645
Report Definition Language 645
Мастер отчетов 646
Сервер отчетов 646
Создание отчета на основе реляционной базы данных 646
Создание отчетов на основе UDM 654
Проектирование отчета Analysis Services 655
Изменение отчетов Analysis Services 660
Пользовательские агрегации 670
Развертывание отчета 674
Управление отчетами Analysis Services 675
Защита данных и выполнение отчета 676
Автоматизация отчетов 679
Создание нерегламентированных отчетов с помощью Microsoft Report Builder 680
Модель отчета 680
Нерегламентированные отчеты 683
Резюме 687

Глава 18. Проектирование кубов, работающих в реальном времени 688

Упреждающее кэширование 689
Упреждающее кэширование на практике 694
Сценарий с длительным временем ожидания 699

Упреждающее кэширование с использованием обновлений по расписанию 702
Сценарий со средним временем ожидания 703
Упреждающее кэширование с параметром хранения MOLAP 704
Сценарий без времени ожидания 708
Параметр хранения Real-time ROLAP 709
Миллиарды и миллиарды записей 710
Резюме 711

Глава 19. Защита данных в Analysis Services 712

Защита исходных данных 713
Защита измерений 715
Сценарий с использованием защиты измерений 716
Защита данных куба 739
Сценарий с использованием защиты ячеек 739
Резюме 748

Приложение. Функции и операторы MDX 749

Основные операторы 749
Операторы, управляющие значениями 749
Формирование кортежей 750
Формирование наборов 751
Функции и операторы 752

Предметный указатель 827