

THE EXPERT'S VOICE® IN SQL SERVER 2005

Foundations of SQL Server 2005 Business Intelligence

*What every SQL Server 2005 user needs
to know to create business intelligence
with SSAS, SSIS, SSRS, and other BI tools*

Lynn Langit

Apress®

Foundations of SQL Server 2005 Business Intelligence



Lynn Langit

Foundations of SQL Server 2005 Business Intelligence

Copyright © 2007 by Lynn Langit

All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval system, without the prior written permission of the copyright owner and the publisher.

ISBN-13 (pbk): 978-1-59059-834-4

ISBN-10 (pbk): 1-59059-834-2

Printed and bound in the United States of America 9 8 7 6 5 4 3 2 1

Trademarked names may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, we use the names only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

Lead Editor: James Huddleston

Technical Reviewer: Matthew Roche

Editorial Board: Steve Anglin, Ewan Buckingham, Gary Cornell, Jason Gilmore, Jonathan Gennick,

Jonathan Hassell, James Huddleston, Chris Mills, Matthew Moodie, Jeff Pepper, Paul Sarknas, Dominic Shakeshaft, Jim Sumser, Matt Wade

Project Manager: Beth Christmas

Copy Edit Manager: Nicole Flores

Copy Editor: Julie McNamee

Assistant Production Director: Kari Brooks-Copony

Production Editor: Kelly Gunther

Composer: Patrick Cunningham

Proofreader: Nancy Sixsmith

Indexer: Carol Burbo

Artist: April Milne

Cover Designer: Kurt Krames

Manufacturing Director: Tom Debolski

Distributed to the book trade worldwide by Springer-Verlag New York, Inc., 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax 201-348-4505, e-mail orders-ny@springer-sbm.com, or visit <http://www.springeronline.com>.

For information on translations, please contact Apress directly at 2560 Ninth Street, Suite 219, Berkeley, CA 94710. Phone 510-549-5930, fax 510-549-5939, e-mail info@apress.com, or visit <http://www.apress.com>.

The information in this book is distributed on an “as is” basis, without warranty. Although every precaution has been taken in the preparation of this work, neither the author(s) nor Apress shall have any liability to any person or entity with respect to any loss or damage caused or alleged to be caused directly or indirectly by the information contained in this work.

Contents at a Glance

About the Author	xiii
About the Technical Reviewer	xv
Acknowledgments	xvii
CHAPTER 1 What Is Business Intelligence?	1
CHAPTER 2 OLAP Modeling	25
CHAPTER 3 Introducing SSIS	51
CHAPTER 4 Using SSAS	73
CHAPTER 5 Intermediate OLAP Modeling	95
CHAPTER 6 Advanced OLAP Modeling	113
CHAPTER 7 Cube Storage and Aggregation	133
CHAPTER 8 Intermediate SSIS	159
CHAPTER 9 Advanced SSIS	197
CHAPTER 10 Introduction to MDX	219
CHAPTER 11 Introduction to Data Mining	243
CHAPTER 12 Reporting Tools	277
CHAPTER 13 SSAS Administration	305
CHAPTER 14 Integration with Office 2007	329
INDEX	369

Contents

About the Author	xiii
About the Technical Reviewer	xv
Acknowledgments	xvii
CHAPTER 1 What Is Business Intelligence?	1
Just What Is BI?	1
Defining BI Using Microsoft's Tools	4
What Microsoft Products Are Involved?	5
BI Languages	8
Understanding BI from an End User's Perspective	10
Demonstrating the Power of BI Using Excel 2003 Pivot Tables	10
Understanding BI Through the Sample	20
Understanding the Business Problems that BI Addresses	22
Reasons to Switch to Microsoft's BI Tools	23
Summary	24
CHAPTER 2 OLAP Modeling	25
Modeling OLAP Source Schemas—Stars	25
Understanding the Star Schema	26
Understanding a Dimension Table	27
Why Create Star Schemas?	30
Effectively Creating Star Schema Models	
Using Grain Statements	32
Tools for Creating Your OLAP Model	33
Modeling Source Schemas—Snowflakes and Other Variations	36
Understanding the Snowflake Schema	36
Knowing When to Use Snowflakes	39
Considering Other Possible Variations	40
Choosing Whether to Use Views Against	
the Relational Data Sources	40

Understanding Dimensional Modeling (UDM)	40
Using the UDM	41
The Slowly Changing Dimension (SCD)	43
The Rapidly Changing Dimension (RCD)	45
Writeback Dimension	45
Understanding Fact (Measure) Modeling	45
Calculated Measure vs. Derived Measure	47
Other Types of Modeling	48
Data Mining	48
KPIs (Key Performance Indicators)	48
Actions, Perspectives, Translations	48
Source Control and Other Documentation Standards	48
Summary	49
CHAPTER 3 Introducing SSIS	51
Understanding ETL	51
Data Maps	53
Staging Servers	55
ETL Tools for BI/SSIS Packages	56
Basic SSIS Packages Using BIDS	59
Developing SSIS Packages	60
Designing SSIS Packages	62
Adding Transformations to the Data Flow	68
Summary	71
CHAPTER 4 Using SSAS	73
Using BIDS to Build a Cube	73
Building Your First Cube	76
Refining Your Cube	84
Reviewing Measures	84
Reviewing Dimensions: Attributes	85
Reviewing Dimensions: Hierarchies	87
Reviewing Dimensions: Member Properties	91
Summary	93
CHAPTER 5 Intermediate OLAP Modeling	95
Adding Key Performance Indicators (KPIs)	95
Implementing KPIs in SSAS	96
Considering Other KPI Issues	100

Using Perspectives and Translations	100
Perspectives	100
Translations	102
Localizing Measure Values	103
Using Actions	107
Other Types of Modeling	112
Summary	112
CHAPTER 6 Advanced OLAP Modeling	113
Multiple Fact Tables in a Single Cube	113
Considering Nulls	117
Modeling Nonstar Dimensions	119
Snowflake Dimensions	119
Degenerate Dimensions	121
Parent-Child Dimensions	121
Many-to-Many Dimensions	123
Role-Playing Dimensions	125
Writeback Dimensions	125
Modeling Changing Dimensions and More	126
Error Handling for Dimension Attribute Loads	127
Using the Business Intelligence Wizard	129
What's Next?	132
Summary	132
CHAPTER 7 Cube Storage and Aggregation	133
Using the Default Storage: MOLAP	133
XMLA (XML for Analysis)	133
Aggregations	135
MOLAP as Default in SSAS	137
Adding Aggregations	137
Advanced Storage: MOLAP, HOLAP, or ROLAP	141
Considering Other Types of Storage	141
ROLAP Dimensions	144
Huge Dimensions	145
Summarizing OLAP Storage Options	146
Using Proactive Caching	147
Notification Settings for Proactive Caching	149
Fine-Tuning Proactive Caching	150

Deciding Among OLTP Partitioning, OLAP Partitioning, or Both	151
Relational Table Partitioning in SQL Server 2005	151
Other OLAP Partition Configurations.	152
Cube and Dimension Processing Options	153
What's Next?	158
Summary	158

CHAPTER 8 Intermediate SSIS	159
General ETL Package-Design Best Practices	159
Creating the SSIS Package from Scratch	160
Configuring Connections.	165
Using Data Source Views (DSVs).	166
Reviewing the Included Samples Packages	167
Adding Control Flow Tasks	168
Container Tasks	170
SQL Tasks	171
File System Tasks	173
Operating System Tasks.	174
Script Tasks	174
Remote Tasks.	175
SSAS Tasks.	175
Precedence Constraints	177
Using Expressions with Precedence Constraints	178
Understanding Data Flow Transformations.	180
Understanding Data Sources and Destinations	180
Adding Transformations to the Data Flow.	182
Adding Data Transformations	184
Split Data Transformations.	185
Translate Data Transformations	187
SSAS Data Transformations.	189
Slowly Changing Dimension Transformation	189
Sample Data Transformations.	192
Run Command Data Transformations	192
Enterprise Edition Only Data Transformations	193
Using the Dynamic Package Configuration Wizard	194
SSIS Expressions	195
Summary	196

CHAPTER 9	Advanced SSIS	197
	Understanding Package Execution	197
	Data Viewers	199
	Debugging SSIS Packages	201
	Logging Execution Results	203
	Error Handling	205
	Event Handlers	207
	Deploying the Package and Configuring Runtime Settings	209
	SSIS Package Deployment Options	209
	SSIS Package Execution Options	211
	SSIS Package Security	214
	Placing Checkpoints	215
	Using Transactions in SSIS Packages	216
	Summary	217
CHAPTER 10	Introduction to MDX	219
	Understanding Basic MDX Query Syntax	219
	Writing Your First MDX Query	224
	Members, Tuples, and Sets	225
	Adding Calculated Members, Named Sets, and Script Commands	226
	Using Calculated Measures	229
	Named Sets	231
	Script Commands	232
	Understanding Common MDX Functions	234
	New or Updated MDX Functions	237
	Adding .NET Assemblies to Your SSAS Project	240
	Configuring Assemblies	241
	Summary	242
CHAPTER 11	Introduction to Data Mining	243
	Defining SSAS Data Mining	243
	More Data Mining Concepts	246
	Architectural Considerations	247
	Reviewing Data Mining Structures	248
	Mining Structure Viewers	252
	Mining Accuracy Charts	255
	Mining Prediction Viewers	256

Understanding the Nine Included Data Mining Algorithms	257
Using the Mining Structure Wizard	265
Content and Data Types	267
Processing Mining Models	271
SSIS and Data Mining	273
Working with the DMX Language	274
A Simple DMX Query	274
Data Mining Clients	275
Summary	276
CHAPTER 12 Reporting Tools	277
Using Excel 2003: Pivot Charts and More	277
Limitations of Excel 2003 as an SSAS Client	282
Using SQL Server Reporting Services (SSRS)	282
Producing Reports with Report Builder	293
Working with .NET 2.0 Report Viewer Controls	298
Understanding SharePoint 2003 Web Parts	300
Examining Business Scorecard Manager (BSM) 2005	302
Considering ProClarity and Data Mining Clients	302
ProClarity	303
Data Mining Clients	303
Summary	304
CHAPTER 13 SSAS Administration	305
Understanding Offline vs. Online Mode in BIDS	305
Reviewing SSMS/SSAS Administration	307
XML for Analysis (XMLA)	308
SSAS Deployment Wizard	310
Server Synchronization	312
Thinking About Disaster Recovery	313
Considering Security	315
Connection Strings	317
Security Roles	318
Other Security Planning Issues	321
Understanding Performance Tuning	321
Applying Scalability	324
Using High Availability Clustering	327
Summary	328