

Pro SQL Server 2012 BI Solutions



Randal Root
Caryn Mason

Apress®

Pro SQL Server 2012 BI Solutions

Copyright © 2012 by Randal Root and Caryn Mason

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

ISBN 978-1-4302-3488-3

ISBN 978-1-4302-3489-0 (eBook)

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

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

President and Publisher: Paul Manning

Lead Editor: James Markham

Technical Reviewers: Maradelyn Taylor-Root and Robert Glancy

Editorial Board: Steve Anglin, Ewan Buckingham, Gary Cornell, Louise Corrigan, Morgan Ertel, Jonathan Gennick, Jonathan Hassell, Robert Hutchinson, Michelle Lowman, James Markham, Matthew Moodie, Jeff Olson, Jeffrey Pepper, Douglas Pundick, Ben Renow-Clarke, Dominic Shakeshaft, Gwenan Spearing, Matt Wade, Tom Welsh

Coordinating Editors: Corbin Collins and Mark Powers

Copy Editor: Kim Wimpsett

Compositor: SPi Global

Indexer: SPi Global

Artist: SPi Global

Cover Designer: Anna Ishchenko

Distributed to the book trade worldwide by Springer Science+Business Media New York, 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 www.springeronline.com.

For information on translations, please e-mail rights@apress.com, or visit www.apress.com.

Apress and friends of ED books may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Special Bulk Sales–eBook Licensing web page at www.apress.com/bulk-sales.

Any source code or other supplementary materials referenced by the author in this text is available to readers at www.apress.com/9781430234883. For detailed information about how to locate your book's source code, go to www.apress.com/source-code.

This book is dedicated to my friends and family. Writing does not come easily to me, and without their support, I doubt that I would have survived the process! Specifically, I would like to thank my coauthor and friend, Caryn, who made this project manageable and more fun than it would have been had I tried doing it solo. Finally, I am especially grateful to Margot Alice for her help as a professional editor. She not only edited this book but also provided encouragement and guidance when my path was uncertain. A journey, however hard, is best shared with friends and family!

—Randal Root

For my two amazing children, Kaylie and Collin, who are my world. May your lives be beautiful, wonderous, and delightfully fulfilling, just as you make my life—every day!

And in honor of Anthony David, who would have been so proud.

—Caryn Mason

Contents at a Glance

About the Authors	xxiii
About the Technical Reviewers	xxv
Acknowledgments	xxvii
■ Chapter 1: Business Intelligence Solutions	1
■ Chapter 2: A Big-Picture Overview	13
■ Chapter 3: Planning Solutions	85
■ Chapter 4: Designing a Data Warehouse	131
■ Chapter 5: Creating a Data Warehouse	161
■ Chapter 6: ETL Processing with SQL	213
■ Chapter 7: Beginning the ETL Process with SSIS	253
■ Chapter 8: Concluding the ETL Process with SSIS	301
■ Chapter 9: Beginning the SSAS Project	335
■ Chapter 10: Configuring Dimensions with SSAS	389
■ Chapter 11: Creating and Configuring SSAS Cubes	435
■ Chapter 12: Additional Cube and Dimension Configurations	483
■ Chapter 13: Creating Reports with SQL Queries	529
■ Chapter 14: Reporting with MDX Queries	565
■ Chapter 15: Reporting with Microsoft Excel	615
■ Chapter 16: Creating Reports with SSRS	653
■ Chapter 17: Configuring Reports with SSRS	693

■ Chapter 18: Testing and Tuning BI Solutions	731
■ Chapter 19: Approve, Release, and Prepare	767
Index	797

Contents

- About the Authors..... xxiii**
- About the Technical Reviewers xxv**
- Acknowledgments xxvii**

- Chapter 1: Business Intelligence Solutions 1**
 - Who Should Read This Book?..... 1
 - What Is a Business Intelligence Solution?..... 2
 - Step 1: Interview and Identify Data 2
 - Step 2: Plan the BI Solution 3
 - Step 3: Create a Data Warehouse 3
 - Step 4: Create an ETL Process..... 4
 - Step 5: Create Cubes 5
 - Step 6: Create Reports..... 6
 - Step 7: Test and Tune the Solution..... 8
 - Step 8: Approve, Release and Prepare..... 8
 - Practice Exercises and More 8
 - Downloadable Content 8
 - Our Example Scenarios 9
 - Setup Instructions 9
 - Think Small, Win Big..... 10
 - Rapid Application Development for BI Solutions 11
 - Moving On 11
 - What’s Next? 12

Chapter 2: A Big-Picture Overview	13
The 10,000-Foot View	13
Interviewing and Isolating Data.....	14
Plan the Solution	15
Creating Planning Documents	16
Adding Documents to Visual Studio.....	17
Creating the Data Warehouse.....	27
An Example Data Warehouse.....	27
Using SQL Code to Create a Data Warehouse.....	28
Running SQL Code from Visual Studio	31
Create the ETL Process	39
ETL with an SSIS Project	39
Creating an SSIS Package	41
Outlining the Control Flow Tasks	43
SSIS Connections	44
Configuring an Execute SQL Task.....	46
Configuring Data Flow Tasks	47
Configuring Additional Data Flows.....	48
Configuring a Data Source	49
Executing an SSIS Task	50
Completing the Package Execution	51
Creating a Cube.....	57
Making a Connection to the Data Warehouse.....	59
Creating a Data Source View	61
Creating Dimensions	64
Creating Cubes	65
Deploying and Processing	66
Creating Reports	74
Using the SSRS Wizard	75
Manually Creating SSRS Reports	81
Testing the Solution.....	84

Approve, Release, and Prepare.....	84
Moving On	84
What’s Next?	84
■ Chapter 3: Planning Solutions	85
Outline the Steps in the Process	86
Interviewing	87
Why Do We Need It?	88
What Are We Building?	89
How Long Will It Take to Build?	91
How Will We Build It?	92
Who Will We Get to Build It?.....	92
When Will We Need It?	94
How Will We Finish It?	94
“Hey, Wait! I’m a Developer, Not a Manager”	95
Documenting the Requirements.....	97
Locating Data	101
Defining the Roles	118
Defining the Team	118
Determining the Schedule.....	119
The IT, Security and Licensing Requirements	121
Estimating the Cost	122
Documenting the Solution Plan	122
Implementation	123
Moving On	129
What’s Next?	129
■ Chapter 4: Designing a Data Warehouse	131
What Is a Data Warehouse?.....	131
What Is a Data Mart?.....	131
Competing Definitions	132

Starting with an OLTP Design	133
A Typical OLTP Database Design	134
Normalized Tables	134
Table Relationships.....	135
Many-to-Many Tables	135
One-to-Many Tables.....	135
Parent–Child One-to-Many Tables	135
A Typical Data Warehouse Database Design	136
Measures	136
The Fact Table.....	138
Dimensions	138
Stars and Snowflakes	139
Performance Considerations	140
Comparing Designs	141
Foreign Keys.....	143
Missing Features	143
Dimensional Patterns	144
Standard Dimensions	144
Fact or Degenerate Dimensions	144
Time Dimensions	145
Role-Playing Dimensions.....	148
Parent–Child Dimensions	149
Junk Dimensions	150
Many-to-Many Dimensions	151
Conformed Dimensions	155
Adding Surrogate Keys	155
Slowly Changing Dimensions	156
Type I	156
Type II	156
Type III	157

Moving On	159
What's Next?	159
■ Chapter 5: Creating a Data Warehouse.....	161
SQL Server Management Studio	161
Connecting to Servers	162
Configuration Manager.....	165
Management Studio Windows	168
Object Explorer	168
The Query Window.....	169
Changing the Query Window Focus.....	170
Executing a Query.....	170
Creating Data Warehouse Database.....	171
Setting the Database Owner	172
Setting the Database Size	173
Setting the Recovery Model.....	175
Keeping Data Warehouse Backups.....	177
Using the Filegroups Option	178
Creating Tables.....	181
Using the Table Designer	181
Using the Diagramming Tool.....	186
Using the Query Window	190
Creating a Date Dimension Table	194
Getting Organized.....	197
Backing Up the Data Warehouse	197
Scripting the Database	197
Organizing Your Files with Visual Studio.....	201
Moving On	210
What's Next?	211
■ Chapter 6: ETL Processing with SQL.....	213
Performing the ETL Programming.....	213

Deciding on Full or Incremental Loading.....	214
Isolating the Data to Be Extracted.....	218
Formatting Your Code	219
Identifying the Transformation Logic.....	219
Programming Your Transformation Logic	220
Reducing the Data	220
Using Column Aliases	221
Converting the Data Types.....	222
Looking Up Surrogate Key Values.....	223
Provide Conformity	224
Generate Date Data	226
Dealing with Nulls	227
The SQL Query Designer	233
Updating Your BI Documentation.....	239
Building an ETL Script	240
Working in the Abstract.....	245
Views	246
Stored Procedures	247
User-Defined Functions	249
Moving On	251
What's Next?	251
■ Chapter 7: Beginning the ETL Process with SSIS	253
Starting Your SSIS Project	254
Adding a Project to an Existing Solution	255
Renaming Your SSIS Package	256
The Anatomy of an SSIS Package	259
The Control Flow Tab	260
The Data Flow Tab	261
Using Sequence Containers.....	265
Using Precedence Constraint Arrows	267

SSIS Variables.....	270
Outlining Your ETL Process.....	273
Data Connections	278
The File Connection Manager.....	279
The OLE DB Connection Manager.....	280
The ADO.NET Connection Manager.....	280
Configuring a Connection	281
Execute SQL Tasks	284
Editing Your Execute SQL Task.....	284
Executing Your Execute SQL Tasks	286
The Progress/Execution Results Tabs	290
Resetting Your Destination Database	292
Moving On	300
What's Next?	300
■ Chapter 8: Concluding the ETL Process with SSIS.....	301
Data Flows.....	301
Outlining a Data Flow Task.....	303
Configuring the Data Source	303
The OLE DB Source Editor	304
The Connection Manager Page.....	304
The Columns Manager Page.....	308
The Error Output Page	309
Data Flow Paths	309
Error Outputs Paths.....	310
Configuring the Data Destination	312
The Connection Manager Page.....	313
Mappings Page.....	315
Error Output Page	316
Executing the Entire Package	330

Moving On	333
What's Next?	334
■ Chapter 9: Beginning the SSAS Project	335
SQL Server vs. Analysis Server Databases.....	335
OLAP Cubes vs. Reporting Tables.....	338
SQL Server vs. Analysis Server Applications.....	339
SSAS Projects.....	341
Data Sources	344
Define a Connection	344
Impersonation Information	345
Data Source Views	351
The Data Source View Wizard.....	352
The Data Source View Designer	357
Dimensions.....	366
The Dimension Wizard	366
Moving On	388
What's Next?	388
■ Chapter 10: Configuring Dimensions with SSAS	389
The Dimension Designer	389
Dimension Structure Tab	390
Attribute Relationships Tab.....	408
Testing Your Progress	411
Translations Tab.....	418
Browser Tab.....	419
Build, Deploy, and Process	419
Building	419
Deploying.....	421
Processing.....	428

Browsing the Dimension	431
Moving On	434
What's Next?	434
■ Chapter 11: Creating and Configuring SSAS Cubes	435
Creating Cubes	435
Processing the Cube	441
Configuring Cubes	443
The Browser Tab	443
Validate the Measures	446
Review the Dimensions	447
Validating the Results	448
The Cube Structure Tab	449
Common Measure Properties	449
Measure Group Properties	452
The Dimension Usage Tab	453
Configuring a Relationship	454
The Calculations Tab	460
Adding a Calculated Member	461
Configuring a Calculated Member	462
Calculated Members vs. Derived Members	464
Making a Test Copy of a Cube	467
KPIs	476
Moving On	481
What's Next?	481
■ Chapter 12: Additional Cube and Dimension Configurations	483
Additional Cube Configurations	483
Actions	483
Partitions	490
Aggregations	501
Perspectives	508

Translations	508
Browser	509
Additional Dimension Configurations	509
Parent-Child Dimensions	510
Role-Playing Dimensions.....	513
Reference Dimensions.....	515
Managing Your Cubes and Dimensions	520
SQL Server Management Studio.....	520
Visual Studio (Live)	522
Moving On	527
What's Next?	528
■ Chapter 13: Creating Reports with SQL Queries	529
Identifying the Data	530
Joining Table Data	531
Ordering Results.....	533
Formatting Results Using SQL Functions	537
Filtering Results	538
Adding Dynamic Filters with Parameters	544
Adding Aggregations	546
Using Subqueries	549
Creating KPI Queries	550
Adding Abstraction Layers.....	552
Using Views	552
Using Stored Procedures	555
Using Your Code in Reporting Applications	561
Moving On	563
What's Next?	564
■ Chapter 14: Reporting with MDX Queries	565
Key Concepts and Terms	565

Programming with MDX	568
Comments	568
Basic and Raw Syntax	569
Running Your MDX Code	569
Optional Syntax.....	571
Default Members.....	573
Using Key vs. Name Identifiers.....	575
Using the Axis 0 and 1 Instead of Column and Row	578
Cells and Tuples.....	579
Calculated Members.....	584
Member Properties	587
Members and Levels	589
The NonEmpty Function.....	590
The Non Empty Clause	592
Member and Level Paths.....	593
Common Functions.....	596
PrevMember and NextMember Functions	596
Using Your Code in Reporting Applications	609
Moving On	613
What Next?	614
■ Chapter 15: Reporting with Microsoft Excel.....	615
Microsoft's BI Reporting.....	615
Excel Reports from the Data Warehouse.....	619
Creating a Connection	619
Creating a Report.....	624
Configuring a Report.....	625
Changing Connection Properties	628
Reconfiguring a Connection	629
Using Stored Procedures.....	632
Working with Excel Reports from a Cube.....	633
Connecting to Your Cube.....	633

Testing Your Reports.....	637
Creating Charts	639
Adding a Header	643
Saving to PDF	644
Moving On	651
What's Next?	652
■ Chapter 16: Creating Reports with SSRS.....	653
SSRS Architecture	653
Developer Tools	654
The Administrative Services	656
SSRS Services	659
SSRS Databases	660
SSRS Configuration Manager	661
Creating SSRS Objects	668
Data Sources	669
Datasets	671
Reports	674
Managing the Report.....	688
Moving On	690
What's Next?	691
■ Chapter 17: Configuring Reports with SSRS.....	693
Creating a Report Template	693
Adding a Header and Footer.....	695
Setting Report Properties	697
Page Size and Report Margins	697
Designing the Header.....	698
Rectangles.....	699
Textboxes.....	700
Images.....	701
Lines	703

Renaming Report Items.....	707
Using Expressions	708
Category and Item Panes.....	710
Expression Editing Pane	711
Placeholders	712
Variables.....	714
Completing the Header.....	716
Configuring the Footer.....	717
Saving the Report Template	721
Using Network Templates.....	721
Using Local Templates.....	722
Moving On	728
What's Next	729
■ Chapter 18: Testing and Tuning BI Solutions.....	731
Testing the BI Solution.....	731
Validation.....	732
Improvement Identification.....	734
Tuning the BI Solution	735
ETL Performance	735
Processing Performance	740
Reporting Performance	741
Common Design Strategies	743
Performance Measurements.....	744
Absolute Performance	744
Relative Performance Measurements	748
Measuring Performance with SQL Profiler	748
An Example Scenario.....	752
Creating a Metadata Database.....	755
Moving On	765
What's Next	766

Chapter 19: Approve, Release, and Prepare	767
The End of the Cycle.....	767
The Final Approval Process	768
The Sign-Off Document	768
Announcing the Release.....	769
The Press Release Title.....	769
The Press Release Body	769
The Press Release Boilerplate.....	770
Releasing the Solution	770
Collecting the Solution Artifacts	771
Deploying the Files	772
Release Documentation	778
SDKs	778
Developer Specifications	782
User Documentation.....	783
Style Guides.....	784
User Manuals.....	784
Help Files	792
User Training	793
Say Thank You	794
Moving On	794
What's Next?	794
Index.....	797

Foreword

I first met Randal as one of his programming students. His teaching style really opened up the subject for me in ways that I had not previously understood. He taught the class from the perspective of someone who had once started from square one himself. And because of this, he was able to convey challenging subjects in a simplified manner. I was the type to ask a million questions, and I began to realize that he was exceedingly knowledgeable about the subjects he taught. It ended up becoming the beginning of a great friendship.

Later he asked me to work with him on projects. He knew that my forté was writing, and we learned that we also worked well together as a team. When he asked me to join him in writing this book, it seemed rather overwhelming to me to even consider it. But he believed in my ability to grasp complicated topics and knew that I would do well in helping translate high-level concepts into readable text. And my need to ask many questions seemed to help the project along as well.

Because the idea for this book began with Randal, I asked him to convey his vision in his own words:

"I've read many books that talk about creating BI projects. They would start with a database project and move onto various topics from there, but all of these books left out fundamental things such as the planning and documentation. I thought there was a need to break it down into more bite-sized pieces and explain the creation of a BI solution in my own words. I felt that having taught for over 10 years and seeing the reaction of students to the various ways of discussing/teaching the subject gave me a unique perspective, because I was able to see my students' reactions. I'm not sure that many authors have had that opportunity.

I think that there is a misconception in the industry that BI solutions are difficult and expensive. I wanted to make sure that people understood that this wasn't the case; BI can be inexpensive and provide immediate benefits, even for very small companies.

I've seen a lot of students struggle with topics such as data warehouse design, ETL processing, or OLAP cubes. I was hoping I could do a better job at making complex concepts seem simple. I've always liked to believe that I had good skills in that regard."

I agree that Randal is excellent at this, because I have seen it for myself. And because of this, I chose to join him in writing this book. It has been quite the undertaking but well worth it. I hope we have been able help a BI developer or two along the way!

Caryn Mason

About the Authors



Randal Root is a senior consultant specializing in .NET programming, SQL Server BI solutions, and technical education. Although he has worked in the industry as a network administrator, DBA, and programmer since the 1980s, for the last 10 years he has focused on providing technical training for businesses and schools such as Microsoft, the University of Washington, and Bellevue College. Randal has now authored two books, *Pro SQL Server 2012 BI Solutions* (Apress) and *A Tester's Guide to .NET Programming* (Apress), and obtained several Microsoft professional certifications including MCSE, MCP+I, MCTS, MCDBA, MCAD, and MCT.

RandalRoot@NorthwestTech.org



Caryn Mason is a content developer and computer programmer with more than 15 years of technical writing experience in a variety of industries, including IT and software. She studied at Bellevue College where she obtained certifications in programming writing and computer programming, specializing in web development. She is especially adept at taking technical concepts and making them understandable by both novices and professionals alike. Caryn has a passion for writing, whether it be technical or fiction, and often finds her inspiration through her two children. In addition to coauthoring *Pro SQL Server BI Solutions 2012* (Apress), Caryn works as an independent technical writing consultant and is working on a young-adult novel.

CarynMason@KeystrokePublications.com

About the Technical Reviewers



Maradelyn Taylor-Root has more than a decade of varied experience in several industries including the insurance, airline, automobile, retail, and financial industries. She has worked as the senior business analyst in the Enrollment Planning Department at Seattle public schools for many years but has recently accepted a position as a business analyst consultant at Microsoft with the Business Excellence Group. Maradelyn's exposure to this group has given her new abilities to see the big picture and ascertain the best way to improve current processes and create new ones using business intelligence tools. In her spare time, she teaches evenings for the University of Washington, Bellevue College, and Cascadia Community College. She has a bachelor's degree in computer science and an MCDBA certification.



Robert Glancy is a product data quality specialist working with Microsoft's SQL Server and Access databases. He is skilled in creating professional BI reports using C#, SSRS, SQL, and VBA. Robert has a bachelor's degree in mathematics from the University of Washington.

Acknowledgments

Wow, I cannot believe that it is finally done! I started writing this book about a year and a half ago, and at that time I figured I could complete it in about six to eight months. Oh, how much I have learned!

Chronologically, I would like to thank Jonathan Gennick, our editorial director at Apress. He helped me get started on this project and provided much guidance along the way.

As time went on, I realized that doing the project on my own was too much and that the project needed an actual technical writer involved. So, I turned to my friend and professional writer, Caryn Mason, who thankfully agreed to coauthor the book with me. Thanks to her, we were able to take what I believed was good content and turn it into meaningful information.

Beginning with the early drafts of the book, we abused the good natures of our technical reviewers. Both Maradelyn Taylor-Root and Robert Glancy were invaluable at proofreading and testing our exercises. We think you will find that their hard work makes the exercises smoother and more enjoyable, which in turn makes the learning process more effective.

As we completed our first drafts, more people at Apress became involved, and I would like to thank them for their time and hard work: Corbin Collins, Tracy Brown Collins, Kim Wimpsett, James Markham, and Mark Powers. All worked hard on the project, but Mark especially made a difference in coordinating the final stages of the book. Thanks, all!

Also, I would like to thank my many students who helped the project by reading our draft chapters and Margot Alice, for her work as our final proof-reader and editor.

—Randal Root

First and foremost I would like to thank Randal Root for bringing me in on this project with him. Your lighthearted nature and supportive personality make you such a joy to work with. I always said you have performed the work of at least five BI professionals in writing this book. That may have been an understatement! You put your heart into what you do, and you do it right. It has been an honor.

Thanks to our technical editors; Maradelyn Taylor-Root jumped in—no holds barred—and worked particularly hard on ensuring this book was done right, and Robert Glancy went beyond the call of duty to work all of the material from start to finish.

I would also like to thank everyone at Apress. Specifically, our editorial director Jonathan Gennick has been a pillar from the start. Your unexpected sense of humor is refreshing. I also want to thank our development editor James Markham, our managing coordinating editor Corbin Collins who oversaw the core of this project, Tracy Brown Collins who handled the times of transition, and Mark Powers who brought the project to a smooth finale.

Thank you to our copy editor, Kim Wimpsett, who made me a little insane trying to decide what to capitalize and which word to use at times. Microsoft can be rather ambiguous, and the challenge was quite the learning experience!

Thank you to Margot Alice for her professional editing. For you I know it was a labor of love, and for that I am so grateful. Thanks to James Mason for the endurance and patience required to see me through this project. And I also wanted to thank Collette Steinwert for being such an inspiration to me and for providing me with a benchmark to attain to.

Thanks to all my wonderful friends and family who continue to believe in me and have shown support on this book. You have kept me going not just on this project but on everything while this book sometimes took center stage. Thanks for all your love and support!

—Caryn Mason