A collection of the most authoritative, detailed, and practical guidelines for data warehousing and business intelligence

Over the past 15 years, Ralph Kimball and his unparalleled team of colleagues have written articles covering more than 250 topics that have come to define the field of data warehousing (DW). The Kimball Group Reader presents the best of these articles, fully updated to reflect their current thinking and practices. Covering the complete lifecycle—including project planning, requirements gathering, dimensional modeling, ETL, and business intelligence (BI) and analytics—the authors provide detailed, technical guidance that expands beyond the toolkit essentials to provide fresh insights and best practices.

The Reader begins with a group of articles that established the core insights and principles of data warehousing. The content that follows provides detailed guidance for every step of the data warehouse lifecycle, with in-depth, practical guidance that you can leverage for your own projects. After covering project planning and business requirements, the authors take a dive into all things dimensional. There are more than 100 articles on dimensional modeling alone. These articles have been fully revised, updated, and expanded where needed to reflect current best practices established by the Kimball Group.
The Kimball Group Reader
About the Authors

Ralph Kimball founded the Kimball Group. Since the mid 1980s, he has been the DW/BI industry’s thought leader on the dimensional approach and has trained more than 10,000 IT professionals. Prior to working at Metaphor and founding Red Brick Systems, Ralph co-invented the Star workstation at Xerox’s Palo Alto Research Center (PARC). Ralph has a Ph.D. in Electrical Engineering from Stanford University.

Margy Ross is President of the Kimball Group. She has focused exclusively on DW/BI solutions since 1982 with an emphasis on business requirements analysis and dimensional modeling. Margy graduated with a B.S. in Industrial Engineering from Northwestern University.
Acknowledgments

First, we want to thank the 28,000 current subscribers to the Kimball Design Tips, and the uncounted numbers of you who have visited the Kimball Group web site to peruse our magazine article archive. This book brings the design tips and articles together in what we hope is a very usable form.

The Kimball Group Reader would not exist without the assistance of our business partners. Kimball Group members Bob Becker, Joy Mundy, and Warren Thornthwaite wrote many of the valuable articles and design tips included in the book. Thanks to Julie Kimball for her insightful comments. Thanks also to former Kimball Group member Bill Schmarzo for his contributions on analytic applications.

Bob Elliott, our executive editor at Wiley Publishing, project editor Sara Shlaer, and the rest of the Wiley team have supported this project with skill, encouragement, and enthusiasm. It has been a pleasure to work with them. We also want to thank Doug Henschen, editor of Intelligent Enterprise, and Julie Langenkamp, editor of DM Review, for allowing us to publish the many articles that appeared in their magazines.

To our families, thank you for your support over the fifteen year span during which we wrote these design tips and articles. Julie Kimball and Scott Ross have been a huge positive force in our professional and personal lives. And, of course, thanks to our children, Sara Kimball Smith, Brian Kimball, and Katie Ross, who have grown into adults over the same time!
Contents at a Glance

Introduction ................................................................. xxi
1 The Reader at a Glance ..................................................... 1
2 Before You Dive In ....................................................... 35
3 Project/Program Planning ............................................... 69
4 Requirements Definition ............................................... 113
5 Data Architecture ......................................................... 133
6 Dimensional Modeling Fundamentals .............................. 179
7 Dimensional Modeling Tasks and Responsibilities ............ 209
8 Fact Table Core Concepts ............................................. 233
9 Dimension Table Core Concepts .................................... 285
10 More Dimension Patterns and Case Studies .................... 333
11 Back Room ETL and Data Quality ................................. 425
12 Technical Architecture Considerations .......................... 513
13 Front Room Business Intelligence Applications .............. 589
14 Maintenance and Growth Considerations .................... 651
   Index of Articles ......................................................... 685
   Index ........................................................................ 693
Contents

Introduction ......................................................... xxii

1 The Reader at a Glance .......................... 1
Setting Up for Success .............................. 1
  1.1 Resist the Urge to Start Coding ................. 1
  1.2 Set Your Boundaries .......................... 4
Tackling DW/BI Design and Development ........ 6
  1.3 Data Wrangling ................................ 6
  1.4 Myth Busters ................................ 8
  1.5 Dividing the World ............................ 10
  1.6 Essential Steps for the Integrated Enterprise Data Warehouse ... 13
  1.7 Drill Down to Ask Why ......................... 22
  1.8 Slowly Changing Dimensions ................. 24
  1.9 Judge Your BI Tool through Your Dimensions ... 28
  1.10 Fact Tables ................................... 30
  1.11 Exploit Your Fact Tables ....................... 32

2 Before You Dive In ................................. 35
Historical Perspective ............................... 35
  2.1 The Database Market Splits ..................... 36
  2.2 Bringing Up Supermarts ......................... 38
Dealing with Demanding Realities ................ 45
  2.3 Brave New Requirements for Data Warehousing ... 46
  2.4 Coping with the Brave New Requirements ....... 50
  2.5 Stirring Things Up ............................ 55
  2.6 Design Constraints and Unavoidable Realities ... 58
  2.7 Two Powerful Ideas ............................ 61
  2.8 Data Warehouse Dining Experience ............. 65
## 3 Project/Program Planning

Professional Responsibilities
- 3.1 Professional Boundaries
- 3.2 An Engineer’s View
- 3.3 Beware the Objection Removers
- 3.4 What Does the Central Team Do?
- 3.5 Avoid DW/BI Isolation
- 3.6 Implementation Analysis Paralysis

Justification and Sponsorship
- 3.7 Habits of Effective Sponsors
- 3.8 TCO Starts with the End User
- 3.9 Better Business Skills for BI and Data Warehouse Professionals

Kimball Methodology
- 3.10 Kimball Lifecycle in a Nutshell
- 3.11 Off the Bench
- 3.12 The Anti-Architect
- 3.13 Think Critically When Applying Best Practices
- 3.14 Eight Guidelines for Low Risk Enterprise Data Warehousing
- 3.15 Relating to Agile Methodologies
- 3.16 Is Agile Enterprise Data Warehousing an Oxymoron?

## 4 Requirements Definition

Gathering Requirements
- 4.1 Alan Alda’s Interviewing Tips for Uncovering Business Requirements
- 4.2 More Business Requirements Gathering Dos and Don’ts
- 4.3 Overcoming Obstacles When Gathering Business Requirements
- 4.4 Surprising Value of Data Profiling

Organizing around Business Processes
- 4.5 Focus on Business Processes, Not Business Departments!
- 4.6 Identifying Business Processes
- 4.7 Business Process Decoder Ring
- 4.8 Relationship between Strategic Business Initiatives and Business Processes

Wrapping Up the Requirements
- 4.9 The Bottom-Up Misnomer
## Contents

5 Data Architecture .................................................. 133  
Making the Case for Dimensional Modeling ......................... 133  
  5.1 Is ER Modeling Hazardous to DSS? ......................... 133  
  5.2 A Dimensional Modeling Manifesto ......................... 137  
  5.3 There Are No Guarantees ................................. 145  
Enterprise Data Warehouse Bus Architecture .................... 148  
  5.4 Divide and Conquer ........................................ 148  
  5.5 The Matrix .................................................. 151  
  5.6 The Matrix: Revisited ...................................... 155  
  5.7 Drill Down into a Detailed Bus Matrix .................... 159  
Integration Instead of Centralization ............................. 161  
  5.8 Integration for Real People ................................ 161  
  5.9 Data Stewardship 101: The First Step to Quality  
  and Consistency ............................................. 165  
  5.10 To Be or Not To Be Centralized ......................... 168  
Contrast with the Corporate Information Factory ................ 171  
  5.11 Differences of Opinion .................................... 171  
  5.12 Don't Support Business Intelligence with a Normalized EDW . 176

6 Dimensional Modeling Fundamentals ................................ 179  
Basics of Dimensional Modeling .................................. 179  
  6.1 Fact Tables and Dimension Tables ......................... 179  
  6.2 Drilling Down, Up, and Across ............................ 183  
  6.3 The Soul of the Data Warehouse, Part One: Drilling Down . 186  
  6.4 The Soul of the Data Warehouse, Part Two: Drilling Across . 189  
  6.5 The Soul of the Data Warehouse, Part Three: Handling Time . 191  
  6.6 Graceful Modifications to Existing Fact and Dimension Tables . 194  
Dos and Don'ts ......................................................... 196  
  6.7 Kimball's Ten Essential Rules of Dimensional Modeling . 196  
  6.8 What Not to Do ............................................... 199  
Myths about Dimensional Modeling ................................ 201  
  6.9 Dangerous Preconceptions ................................. 201  
  6.10 Fables and Facts ........................................... 204

7 Dimensional Modeling Tasks and Responsibilities ............ 209  
Design Activities ...................................................... 209  
  7.1 Letting the Users Sleep .................................... 209