

Contents at a Glance

Introduction	1
Part I: The Data Warehouse: Home for Your Data Assets... 7	
Chapter 1: What's in a Data Warehouse?.....	9
Chapter 2: What Should You Expect from Your Data Warehouse?.....	25
Chapter 3: Have It Your Way: The Structure of a Data Warehouse.....	37
Chapter 4: Data Marts: Your Retail Data Outlet.....	59
Part II: Data Warehousing Technology	71
Chapter 5: Relational Databases and Data Warehousing	73
Chapter 6: Specialty Databases and Data Warehousing.....	85
Chapter 7: Stuck in the Middle with You: Data Warehousing Middleware	95
Part III: Business Intelligence and Data Warehousing... 113	
Chapter 8: An Intelligent Look at Business Intelligence.....	115
Chapter 9: Simple Database Querying and Reporting	125
Chapter 10: Business Analysis (OLAP)	135
Chapter 11: Data Mining: Hi-Ho, Hi-Ho, It's Off to Mine We Go.....	149
Chapter 12: Dashboards and Scorecards	155
Part IV: Data Warehousing Projects: How to Do Them Right	163
Chapter 13: Data Warehousing and Other IT Projects: The Same but Different ...	165
Chapter 14: Building a Winning Data Warehousing Project Team	179
Chapter 15: You Need What? When? — Capturing Requirements	193
Chapter 16: Analyzing Data Sources.....	203
Chapter 17: Delivering the Goods.....	213
Chapter 18: User Testing, Feedback, and Acceptance.....	225
Part V: Data Warehousing: The Big Picture.....	231
Chapter 19: The Information Value Chain: Connecting Internal and External Data.....	233
Chapter 20: Data Warehousing Driving Quality and Integration	247
Chapter 21: The View from the Executive Boardroom	263

Chapter 22: Existing Sort-of Data Warehouses: Upgrade or Replace?	271
Chapter 23: Surviving in the Computer Industry (and Handling Vendors).....	281
Chapter 24: Working with Data Warehousing Consultants	291
<i>Part VI: Data Warehousing in the Not-Too-Distant Future</i>	297
Chapter 25: Expanding Your Data Warehouse with Unstructured Data.....	299
Chapter 26: Agreeing to Disagree about Semantics	305
Chapter 27: Collaborative Business Intelligence	311
<i>Part VII: The Part of Tens</i>	317
Chapter 28: Ten Questions to Consider When You're Selecting User Tools	319
Chapter 29: Ten Secrets to Managing Your Project Successfully.....	325
Chapter 30: Ten Sources of Up-to-Date Information about Data Warehousing	331
Chapter 31: Ten Mandatory Skills for a Data Warehousing Consultant	335
Chapter 32: Ten Signs of a Data Warehousing Project in Trouble	339
Chapter 33: Ten Signs of a Successful Data Warehousing Project	343
Chapter 34: Ten Subject Areas to Cover with Product Vendors	347
<i>Index</i>	351

Table of Contents

.....

<i>Introduction</i>	1
Why I Wrote This Book.....	1
How to Use This Book.....	2
Part I: The Data Warehouse: Home for Your Data Assets	3
Part II: Data Warehousing Technology.....	3
Part III: Business Intelligence and Data Warehousing.....	4
Part IV: Data Warehousing Projects: How to Do Them Right.....	4
Part V: Data Warehousing: The Big Picture.....	4
Part VI: Data Warehousing in the Not-Too-Distant Future	5
Part VII: The Part of Tens.....	6
Icons Used in This Book	6
About the Product References in This Book.....	6

Part I: The Data Warehouse: Home for Your Data Assets... 7

Chapter 1: What's in a Data Warehouse?	9
The Data Warehouse: A Place for Your Data Assets.....	9
Classifying data: What is a data asset?.....	10
Manufacturing data assets.....	10
Data Warehousing: A Working Definition	12
Today's data warehousing defined.....	13
A broader, forward looking definition.....	13
A Brief History of Data Warehousing	14
Before our time — the foundation.....	14
The 1970s — the preparation.....	15
The 1980s — the birth.....	16
The 1990s — the adolescent	17
The 2000s — the adult.....	18
Is a Bigger Data Warehouse a Better Data Warehouse?	19
Realizing That a Data Warehouse (Usually) Has a Historical Perspective	20
It's Data Warehouse, Not Data Dump.....	21
Chapter 2: What Should You Expect from Your Data Warehouse?.. .	25
Using the Data Warehouse to Make Better Business Decisions	25
Finding Data at Your Fingertips	28
Facilitating Communications with Data Warehousing.....	30
IT-to-business organization communications	31
Communications across business organizations	32
Facilitating Business Change with Data Warehousing.....	34

Chapter 3: Have It Your Way: The Structure of a Data Warehouse. . . 37

- Ensuring That Your Implementations Are Unique 37
- Classifying the Data Warehouse 38
 - The data warehouse lite..... 41
 - The data warehouse deluxe 46
 - The data warehouse supreme..... 52
- To Centralize or Distribute, That Is the Question 56

Chapter 4: Data Marts: Your Retail Data Outlet 59

- Architectural Approaches to Data Marts 59
 - Data marts sourced by a data warehouse 60
 - Top-down, quick-strike data marts..... 62
 - Bottom-up, integration-oriented data marts 63
- What to Put in a Data Mart 64
 - Geography-bounded data 64
 - Organization-bounded data..... 65
 - Function-bounded data..... 66
 - Market-bounded data 67
 - Answers to specific business questions 67
 - Anything!..... 68
- Data mart or data warehouse?..... 68
- Implementing a Data Mart — Quickly 69

Part II: Data Warehousing Technology 71

Chapter 5: Relational Databases and Data Warehousing 73

- The Old Way of Thinking 73
 - A technology-based discussion: The roots of relational database technology 74
 - The OLAP-only fallacy 77
- The New Way of Thinking..... 78
 - Fine-tuning databases for data warehousing 78
 - Optimizing data access 79
 - Avoiding scanning unnecessary data..... 79
 - Handling large data volume..... 80
- Designing Your Relational Database for Data Warehouse Usage..... 81
 - Looking at why traditional relational design techniques don't work well..... 81
 - Exploring new ways to design a relational-based data warehouse 82
- Relational Products and Data Warehousing 83
 - IBM Data Management family..... 83
 - Microsoft SQL Server..... 84
 - Oracle 84



Chapter 6: Specialty Databases and Data Warehousing 85

- Multidimensional Databases 86
 - The idea behind multidimensional databases 86
 - Are multidimensional databases still worth looking at?..... 90
- Horizontal versus Vertical Data Storage Management 90
- Data Warehouse Appliances 92
- Data Warehousing Specialty Database Products 93
 - Cognos (An IBM company)..... 93
 - Microsoft..... 93
 - Oracle 94
 - Sybase IQ..... 94
 - Vertica 94

**Chapter 7: Stuck in the Middle with You:
Data Warehousing Middleware 95**

- What Is Middleware? 95
- Middleware for Data Warehousing..... 96
 - The services 96
 - Should you use tools or custom code? 98
- What Each Middleware Service Does for You..... 98
 - Data selection and extractions..... 99
 - Data quality assurance, part I 99
 - Data movement, part I..... 101
 - Data mapping and transformation..... 102
 - Data quality assurance, part II 103
 - Data movement, part II..... 104
 - Data loading..... 104
- Specialty Middleware Services 104
 - Replication services for data warehousing 105
 - Enterprise Information Integration services 106
- Vendors with Middleware Products for Data Warehousing 110
 - Composite Software..... 110
 - IBM..... 110
 - Informatica..... 111
 - Ipedo..... 111
 - Microsoft..... 111
 - Oracle 111
 - Sybase (Avaki)..... 112

Part III: Business Intelligence and Data Warehousing ... 113

Chapter 8: An Intelligent Look at Business Intelligence 115

- The Main Categories of Business Intelligence 116
 - Querying and reporting..... 116
 - Business analysis (OLAP) 117
 - Data mining..... 118
 - Dashboards and scorecards..... 119

Other Types of Business Intelligence	120
Statistical processing	121
Geographical information systems.....	121
Mash-ups.....	122
Business intelligence applications	122
Business Intelligence Architecture and Data Warehousing.....	123
Chapter 9: Simple Database Querying and Reporting.....	125
What Functionality Does a Querying and Reporting Tool Provide?	126
The role of SQL.....	127
Technical query tools.....	128
User query tools.....	129
Reporting tools.....	129
The idea of managed queries and reports.....	129
Is This All You Need?	130
Designing a Relational Database for Querying and Reporting Support.....	131
Vendors with Querying and Reporting Products for Data Warehousing.....	133
Business Objects (SAP).....	133
Cognos (IBM).....	133
Information Builders.....	134
Microsoft.....	134
Oracle	134
Chapter 10: Business Analysis (OLAP).....	135
What Is Business Analysis?	136
The OLAP Acronym Parade.....	137
Business analysis (Visualization)	137
OLAP middleware	138
OLAP databases	138
First, an Editorial	139
Business Analysis (OLAP) Features: An Overview.....	139
Drill-down.....	140
Drill-up.....	143
Drill-across.....	143
Drill-through	144
Pivoting	144
Trending.....	145
Nesting	145
Visualizing.....	145
Data Warehousing Business Analysis Vendors	146
IBM.....	146
MicroStrategy	147
Oracle	147
Pentaho	147
SAP	147
SAS	148

Chapter 11: Data Mining: Hi-Ho, Hi-Ho, It's Off to Mine We Go . . . 149

Data Mining in Specific Business Missions.....	150
Data Mining and Artificial Intelligence.....	150
Data Mining and Statistics.....	151
Some Vendors with Data Mining Products.....	152
Microsoft.....	152
SAS.....	152
SPSS.....	153

Chapter 12: Dashboards and Scorecards. 155

Dashboard and Scorecard Principles.....	155
Dashboards.....	156
Scorecards.....	157
The Relationship between Dashboards, Scorecards, and the Other Parts of Business Intelligence.....	158
EIS and Key Indicators.....	158
The Briefing Book.....	159
The Portal Command Center.....	160
Who Produces EIS Products.....	161

***Part IV: Data Warehousing Projects:
How to Do Them Right 163*****Chapter 13: Data Warehousing and Other IT Projects:
The Same but Different 165**

Why a Data Warehousing Project Is (Almost) Like Any Other Development Project.....	166
How to Apply Your Company's Best Development Practices to Your Project.....	167
How to Handle the Uniqueness of Data Warehousing.....	170
Why Your Data Warehousing Project Must Have Top-Level Buy-In.....	174
How Do I Conduct a Large, Enterprise-Scale Data Warehousing Initiative?.....	175
Top-down.....	176
Bottom-up.....	177
Mixed-mode.....	177

**Chapter 14: Building a Winning Data
Warehousing Project Team 179**

Don't Make This Mistake!.....	180
The Roles You Have to Fill on Your Project.....	180
Project manager.....	181
Technical leader.....	183
Chief architect.....	184
Business requirements analyst.....	184

Data modeler and conceptual/logical database designer	185
Database administrator and physical database designer	187
Front-end tools specialist and developer	187
Middleware specialist	188
Quality assurance (QA) specialist	188
Source data analyst	189
User community interaction manager	189
Technical executive sponsor	189
User community executive sponsor	190
And Now, the People	190
Organizational Operating Model	191

Chapter 15: You Need What? When? — Capturing Requirements . . . 193

Choosing between Being Business or Technically Driven	193
Technically-Driven Data Warehousing	194
Subject area	194
Enterprise data modeling	195
Business-Driven Business Intelligence	195
Starting with business questions	197
Accessing the value of the information	198
Defining key business objects	199
Building a business model	201
Prototyping and iterating with the users	201
Signing off on scope	202

Chapter 16: Analyzing Data Sources 203

Begin with Source Data Structures, but Don't Stop There	205
Identify What Data You Need to Analyze	206
Line Up the Help You'll Need	208
Techniques for Analyzing Data Sources and Their Content	209
Analyze What's Not There: Data Gap Analysis	210
Determine Mapping and Transformation Logic	211

Chapter 17: Delivering the Goods 213

Exploring Architecture Principles	213
What's an architecture?	214
What's an adaptable architecture?	214
Understanding Data Warehousing Architectural Keys	215
People and their roles	215
Consistent delivery process	216
Standard delivery platform	216
Assessing Your Data Warehouse Architecture	217
What are you building?	218
How are you building it?	219
Is the delivery automated?	221
Architecting through Abstraction	222

Chapter 18: User Testing, Feedback, and Acceptance 225

Getting Users Involved Early in Data Warehousing	226
Using Real Business Situations	227
Ensuring That Users Provide Necessary Feedback.....	228
After the Scope: Involving Users during Design and Development	229
Understanding What Determines User Acceptance.....	229

Part V: Data Warehousing: The Big Picture 231**Chapter 19: The Information Value Chain:
Connecting Internal and External Data 233**

Identifying Data You Need from Other People	233
Recognizing Why External Data Is Important	234
Viewing External Data from a User's Perspective	235
Determining What External Data You Really Need	236
Ensuring the Quality of Incoming External Data	238
Filtering and Reorganizing Data after It Arrives	240
Restocking Your External Data	240
Acquiring External Data.....	242
Finding external information.....	242
Gathering general information.....	243
Cruising the Internet.....	243
Maintaining Control over External Data	243
Staying on top of changes.....	244
Knowing what to do with historical external data	244
Determining when new external data sources are available.....	245
Switching from one external data provider to another	245

Chapter 20: Data Warehousing Driving Quality and Integration . . . 247

The Infrastructure Challenge	248
Data Warehouse Data Stores.....	249
Source data feeds.....	250
Operational data store (ODS).....	250
Master data management (MDM)	258
Service-oriented architecture (SOA)	259
Dealing with Conflict: Special Challenges to Your Data Warehousing Environment	260

Chapter 21: The View from the Executive Boardroom. 263

What Does Top Management Need to Know?.....	264
Tell them this	265
Keep selling the data warehousing project	266
Data Warehousing and the Business-Trends Bandwagon.....	267
Data Warehousing in a Cross-Company Setting	268
Connecting the Enterprise.....	270

Chapter 22: Existing Sort-of Data Warehouses: Upgrade or Replace? 271

- The Data Haves and Have-Nots..... 272
 - The first step: Cataloguing the extract files, who uses them, and why..... 274
 - And then, the review 276
- Decisions, Decisions..... 276
 - Choice 1: Get rid of it..... 277
 - Choice 2: Replace it 277
 - Choice 3: Retain it 278
- Caution: Migration Isn't Development — It's Much More Difficult 279
- Beware: Don't Take Away Valued Functionality..... 280

Chapter 23: Surviving in the Computer Industry (and Handling Vendors). 281

- How to Be a Smart Shopper at Data Warehousing
 - Conferences and Trade Shows 283
 - Do your homework first 284
 - Ask a lot of questions 284
 - Be skeptical..... 285
 - Don't get rushed into a purchase 285
 - Dealing with Data Warehousing Product Vendors..... 286
 - Check out the product and the company before you begin discussions 286
 - Take the lead during the meeting..... 287
 - Be skeptical — again 288
 - Be a cautious buyer 288
- A Look Ahead: Data Warehousing, Mainstream Technologies, and Vendors..... 289

Chapter 24: Working with Data Warehousing Consultants 291

- Do You Really Need Consultants to Help Build a Data Warehouse?..... 291
- Watch Out, Though! 292
- A Final Word about Data Warehousing Consultants..... 295

Part VI: Data Warehousing in the Not-Too-Distant Future 297

Chapter 25: Expanding Your Data Warehouse with Unstructured Data 299

- Traditional Data Warehousing Means Analyzing
 - Traditional Data Types..... 299
 - It's a Multimedia World, After All. 300

How Does Business Intelligence Work with Unstructured Data?	301
An Alternative Path: From Unstructured Information to Structured Data	303
Chapter 26: Agreeing to Disagree about Semantics	305
Defining Semantics	305
Emergence of the Semantic Web?	306
Preparing for Semantic Data Warehousing	307
Starting Out on Your Semantic Journey	308
Business intelligence semantic layer management	309
Business rules management	309
Chapter 27: Collaborative Business Intelligence	311
Future Business Intelligence Support Model	312
Knowledge retention	313
Knowledge discovery	313
Knowledge proliferation	313
Leveraging Examples from Highly Successful Collaboration Solutions.....	314
Rate a report.....	314
Report relationships.....	314
Find a report	314
Find the meaning.....	315
Shared interests — shared information.....	315
Visualization	315
The Vision of Collaborative Business Intelligence	316
 Part VII: The Part of Tens.....	 317
Chapter 28: Ten Questions to Consider When You're Selecting User Tools	319
Do I Want a Smorgasbord or a Sit-Down Restaurant?	319
Can a User Stop a Runaway Query or Report?	320
How Does Performance Differ with Varying Amounts of Data?.....	321
Can Users Access Different Databases?.....	322
Can Data Definitions Be Easily Changed?	322
How Does the Tool Deploy?	322
How Does Performance Change If You Have a Large Number of Users?.....	323
What Online Help and Assistance Is Available, and How Good Is It? ...	323
Does the Tool Support Interfaces to Other Products?	324
What Happens When You Pull the Plug?	324

Chapter 29: Ten Secrets to Managing Your Project Successfully	325
Tell It Like It Is.....	325
Put the Right People in the Right Roles.....	326
Be a Tough but Fair Negotiator	326
Deal Carefully with Product Vendors	326
Watch the Project Plan	327
Don't Micromanage	327
Use a Project Wiki.....	327
Don't Overlook the Effect of Organizational Culture	328
Don't Forget about Deployment and Operations	329
Take a Breather Occasionally	329
Chapter 30: Ten Sources of Up-to-Date Information about Data Warehousing	331
The Data Warehousing Institute.....	331
The Data Warehousing Information Center	332
The OLAP Report.....	332
Intelligent Enterprise.....	332
b-eye Business Intelligence Network	333
Wikipedia.....	333
DMReview.com	333
BusinessIntelligence.com	333
Industry Analysts' Web Sites	334
Product Vendors' Web Sites	334
Chapter 31: Ten Mandatory Skills for a Data Warehousing Consultant	335
Broad Vision.....	335
Deep Technical Expertise in One or Two Areas	336
Communications Skills.....	336
The Ability to Analyze Data Sources.....	336
The Ability to Distinguish between Requirements and Wishes	337
Conflict-Resolution Skills	337
An Early-Warning System	337
General Systems and Application Development Knowledge	338
The Know-How to Find Up-to-Date Information.....	338
A Hype-Free Vocabulary	338
Chapter 32: Ten Signs of a Data Warehousing Project in Trouble	339
The Project's Scope Phase Ends with No General Consensus	339
The Mission Statement Gets Questioned after the Scope Phase Ends	340

Tools Are Selected without Adequate Research	340
People Get Pulled from Your Team for “Just a Few Days”	340
You’re Overruled When You Attempt to Handle Scope Creep.....	341
Your Executive Sponsor Leaves the Company	341
You Overhear, “This Will Never Work, but I’m Not Saying Anything”	341
You Find a Major “Uh-Oh” in One of the Products You’re Using	342
The IT Organization Responsible for Supporting the Project Pulls Its Support.....	342
Resignations Begin	342

**Chapter 33: Ten Signs of a Successful
Data Warehousing Project 343**

The Executive Sponsor Says, “This Thing Works — It Really Works!”	343
You Receive a Flood of Suggested Enhancements and Additional Capabilities	344
User Group Meetings Are Almost Full.....	344
The User Base Keeps Growing and Growing and Growing	344
The Executive Sponsor Cheerfully Volunteers Your Company as a Reference Site.....	345
The Company CEO Asks, “How Can I Get One of Those Things?”	345
The Response to Your Next Funding Request Is, “Whatever You Need — It’s Yours.”	345
You Get Promoted — and So Do Some of Your Team Members	346
You Achieve Celebrity Status in the Company	346
You Get Your Picture on the Cover of the Rolling Stone.....	346

Chapter 34: Ten Subject Areas to Cover with Product Vendors . . . 347

Product’s Chief Architect	347
Development Team	348
Customer Feedback.....	348
Employee Retention	348
Marketplace.....	349
Product Uniqueness	349
Clients	349
The Future	350
Internet and Internet Integration Approach	350
Integrity.....	350

***Index*** **351**