

# Data Warehousing Fundamentals for IT Professionals, 2nd Edition

[Paulraj Ponniah](#)

ISBN: 978-0-470-46207-2

## TABLE OF CONTENTS

Preface xxv

### **Part 1 Overview and Concepts 1**

#### **1 The Compelling Need For Data Warehousing 3**

Chapter Objectives 3

Escalating Need For Strategic Information 4

Failures of Past Decision-Support Systems 9

Operational Versus Decision-Support Systems 11

Data Warehousing—The Only Viable Solution 13

A Simple Concept For Information Delivery 15

The Data Warehousing Movement 17

Evolution of Business Intelligence 18

Chapter Summary 20

Review Questions 20

Exercises 21

#### **2 Data Warehouse: The Building Blocks 23**

Chapter Objectives 23

Defining Features 24

Data Warehouses and Data Marts 29

Architectural Types 32

Overview of The Components 34

Metadata in The Data Warehouse 41

Chapter Summary 42

Review Questions 43

Exercises 43

### **3 Trends in Data Warehousing 45**

Chapter Objectives 45

Continued Growth in Data Warehousing 46

Significant Trends 50

Emergence of Standards 64

Web-Enabled Data Warehouse 66

Chapter Summary 69

Review Questions 69

Exercises 70

## **Part 2 Planning and Requirements 71**

### **4 Planning and Project Management 73**

Chapter Objectives 73

Planning Your Data Warehouse 74

The Data Warehouse Project 79

The Development Phases 83

The Project Team 85

Project Management Considerations 90

Chapter Summary 96

Review Questions 96

Exercises 97

### **5 Defining The Business Requirements 99**

Chapter Objectives 99

Dimensional Analysis 100

Information Packages—A Useful Concept 103

Requirements Gathering Methods 109

Requirements Definition: Scope and Content 116

Chapter Summary 119

Review Questions 119

Exercises 120

## **6 Requirements As The Driving Force For Data Warehousing 121**

Chapter Objectives 121

Data Design 122

The Architectural Plan 125

Data Storage Specifications 131

DBMS Selection 132

Information Delivery Strategy 133

Chapter Summary 136

Review Questions 136

Exercises 137

## **Part 3 Architecture and Infrastructure 139**

### **7 Architectural Components 141**

Chapter Objectives 141

Understanding Data Warehouse Architecture 141

Distinguishing Characteristics 143

Architectural Framework 146

Technical Architecture 148

Architectural Types 156

Chapter Summary 160

Review Questions 160

Exercises 161

### **8 Infrastructure As The Foundation For Data Warehousing 163**

Chapter Objectives 163

Infrastructure Supporting Architecture 164

Hardware and Operating Systems 166

Database Software 181

Collection of Tools 184

Data Warehouse Appliances 188

Chapter Summary 191

Review Questions 191

Exercises 192

## **9 The Significant Role of Metadata 193**

Chapter Objectives 193

Why Metadata Is Important 193

Metadata Types By Functional Areas 203

Business Metadata 207

Technical Metadata 209

How To Provide Metadata 212

Chapter Summary 219

Review Questions 220

Exercises 220

## **Part 4 Data Design and Data Preparation 223**

### **10 Principles of Dimensional Modeling 225**

Chapter Objectives 225

From Requirements To Data Design 225

The Star Schema 232

Star Schema Keys 239

Advantages of The Star Schema 241

Star Schema: Examples 244

Chapter Summary 246

Review Questions 247

Exercises 247

### **11 Dimensional Modeling: Advanced Topics 249**

Chapter Objectives 249

Updates To The Dimension Tables 250

Miscellaneous Dimensions 255

The Snowflake Schema 259

Aggregate Fact Tables 262

Families of Stars 272

Chapter Summary 277

Review Questions 278

Exercises 278

## **12 Data Extraction, Transformation, and Loading 281**

Chapter Objectives 281

ETL Overview 282

ETL Requirements and Steps 284

Data Extraction 286

Data Transformation 295

Data Loading 302

ETL Summary 308

Other Integration Approaches 311

Chapter Summary 313

Review Questions 313

Exercises 314

## **13 Data Quality: A Key To Success 315**

Chapter Objectives 315

Why Is Data Quality Critical? 316

Data Quality Challenges 323

Data Quality Tools 326

Data Quality Initiative 328

Master Data Management (Mdm) 335

MDM Categories 335

MDM Benefits 335

MDM and Data Warehousing 336

Chapter Summary 336

Review Questions 336

Exercises 337

## **Part 5 Information Access and Delivery 339**

### **14 Matching Information To The Classes of Users 341**

Chapter Objectives 341

Information From The Data Warehouse 342

Who Will Use The Information? 349

Information Delivery 356

Information Delivery Tools 360

Information Delivery: Special Topics 366

Chapter Summary 371

Review Questions 371

Exercises 372

### **15 OLAP in The Data Warehouse 373**

Chapter Objectives 373

Demand For Online Analytical Processing 374

OLAP is The Answer 379

OLAP Definitions and Rules 379

OLAP Characteristics 382

Major Features and Functions 382

OLAP Models 393

OLAP Implementation Considerations 398

Chapter Summary 404

Review Questions 405

Exercises 405

## **16 Data Warehousing and The Web 407**

Chapter Objectives 407

Web-Enabled Data Warehouse 408

Web-Based Information Delivery 414

OLAP and The Web 420

Building A Web-Enabled Data Warehouse 421

Chapter Summary 426

Review Questions 426

Exercises 427

## **17 Data Mining Basics 429**

Chapter Objectives 429

What Is Data Mining? 430

Major Data Mining Techniques 439

Data Mining Applications 452

Chapter Summary 459

Review Questions 459

Exercises 460

## **Part 6 Implementation and Maintenance 461**

### **18 The Physical Design Process 463**

Chapter Objectives 463

Physical Design Steps 464

Physical Design Considerations 467

Physical Storage 473

Indexing The Data Warehouse 477

Performance Enhancement Techniques 483

Chapter Summary 486

Review Questions 486

Exercises 487

## **19 Data Warehouse Deployment 489**

Chapter Objectives 489

Data Warehouse Testing 490

Major Deployment Activities 491

Considerations For A Pilot 497

Security 502

Backup and Recovery 504

Chapter Summary 508

Review Questions 508

Exercises 509

## **20 Growth and Maintenance 511**

Chapter Objectives 511

Monitoring The Data Warehouse 512

User Training and Support 515

Managing The Data Warehouse 520

Chapter Summary 524

Review Questions 525

Exercises 525

Answers To Selected Exercises 527

Appendix A: Project Life Cycle Steps and Checklists 531

Appendix B: Critical Factors For Success 535

Appendix C: Guidelines For Evaluating Vendor Solutions 537

Appendix D: Highlights of Vendors and Products 539

Appendix E: Real-World Examples of Best Practices 549

References 555

Glossary 557

Index 565



