The Data Warehouse Lifecycle Toolkit Expert Methods For Designing Developing And Deploying Data Warehouses

This book constitutes the refereed proceedings of the 6th International Conference on Data Warehousing and Knowledge Discovery, DaWaK 2004, held in Zaragoza, Spain, in September 2004. The 40 revised full papers presented were carefully reviewed and selected from over 100 submissions. The papers are organized in topical sections on data warehouse design; knowledge discovery framework and XML data mining, data cubes and queries; multidimensional schema and data aggregation; inductive databases and temporal rules; industrial applications; data clustering; data visualization and exploration; data classification, extraction, and interpretation; data semantics, association rule mining; event sequence mining; and pattern mining.

Completely revised, expanded, and updated, this second edition gives extensive new coverage of data integration, management, indexing, cleansing, and transformation. The book covers powerful new multi-dimensional front-ends and conversion tools and gives detailed coverage of lifecycle issues.

Ralph Kimball's three data warehousing books, The Data Warehouse Toolkit, The Data Warehouse Lifecycle Toolkit, and The Data Webhouse Toolkit, provide you with everything you will need to create, manage, and use your data warehouse. His first book, The Data Warehouse Toolkit, is the definitive guide to building a data warehouse. Kimball uses actual case studies of existing data warehouses developed for specific types of business applications such as retail, manufacturing, banking, insurance, subcriptions and airline reservations. Using the techniques learned in Kimball's first book, The Data Warehouse Lifecycle Toolkit carries them to the larger issues of delivering complete data marts and data warehouses. The book shows you all the practical details involved in planning, designing, developing, deploying, and growing data warehouses. The Data Webhouse Toolkit is a groundbreaking guide which introduces the Webhouse, a powerful new way of capturing valuable information flowing into a Web site and ordering it in ways that are useful to managers, strategic decision-makers, and customers.

This is the first book to provide in-depth coverage of star schema aggregates used in dimensional modeling-from selection and design, to loading and usage, to specific tasks and deliverables for implementation projects Covers the principles of aggregate schema design and the pros and cons of various types of commercial solutions for navigating and building aggregates Discusses how to include aggregates in data warehouse development projects that focus on incremental development, iterative builds, and early data loads

Mastering Data Warehouse Design

Data Warehouse Design Solutions

With SQL ServerÂ 2005 and the Microsoft Business Intelligence Toolset

Data Pipelines Pocket Reference

Kimball's Data Warehouse Toolkit Classics

The Analytics Lifecycle Toolkit

A cutting-edge response to Ralph Kimball's challenge to thedata warehouse community that answers some tough questions aboutthe effectiveness of the relational approach to datawarehousing Written by one of the best-known exponents of the Bill Inmonapproach to data warehousing Addresses head-on the tough issues raised by Kimball andexplains how to choose the best modeling technique for solvingcommon data warehouse design problems Weighs the pros and cons of relational vs. dimensional modelingtechniques Focuses on tough modeling problems, including creating andmaintaining keys and modeling calendars, hierarchies, transactions,and data quality

Reduce operating and maintenance costs while substantially improving the performance of new and existing data warehouses and data marts Data Warehouse Performance This book tells you what you need to know to design, build, and manage data warehouses and data marts for optimum performance. Written by an all-star team of data warehouse pioneers and innovators-including Bill Inmon, "the father of the data warehouse," and Ken Rudin, one of the leading experts on performance-the book describes the layers of a high-performance data warehouse environment and guides the reader through their implementation and management. It also supplies proven techniques for supercharging the performance of existing environments. Crucial topics covered include: * Mitigating the impact of dormant data on performance * Data cleansing and implementation techniques * Implementing platform components like data marts to support scalability * Database design, sizing, and optimization techniques, including star schema and indexing * Hardware assessment, selection, and sizing * The role of monitors in balancing workload and assessing performance * Creating a service management contract to meet user expectations

The objective of the workshops associated with the ER'99 18th International Conference on Conceptual Modeling is to give participants access to high level presentations on specialized, hot, or emerging scientific topics. Three themes have been selected in this respect: — Evolution and Change in Data Management (ECDM'99) dealing with han dling the evolution of data and data structure, — Reverse Engineering in Information Systems (REIS'99) aimed at exploring the issues raised by legacy systems, — The World Wide Web and Conceptual Modehng (WWWCM'99) which ana lyzes the mutual contribution of WWW resources and techniques with con ceptual modeling. ER'99 has been organized so that there is no overlap between conference ses sions and the workshops. Therefore participants can follow both the conference and the workshop presentations they are interested in. I would like to thank the ER'99 program co-chairs, Jacky Akoka and Mokrane Bouzeghoub for having given me the opportunity to organize these workshops. I would also like to thank Stephen Liddle for his valuable help in managing the evaluation procedure for submitted papers and helping to prepare the workshop proceedings for publication. August 1999 Jacques Kouloumdjian Preface for ECDM'99 The first part of this volume contains the proceedings of the First International Workshop on Evolution and Change in Data Management, ECDM'99, which was held in conjunction with the 18th International Conference on Conceptual Modehng (ER'99) in Paris, France, November 15-18, 1999.

The first, step-by-step guide to building Web-enabled data warehouses The Web can be an incredibly rich source of customer data, and right now companies across industry sectors are hustling to get up and running with data warehouses capable of capturing the clickstream data from their Web sites. This allows companies to track exactly where a customer is going, or "clicking to," on their site in order to gain meaningful information about that customer's preferences. Following Ralph Kimball's The Data Webhouse Toolkit (0-471-37680-9) where he provides the blueprint, Clickstream Data Warehousing fills developers in on all the technical details that go into building a Web-enabled data warehouse. The authors review all key architectural and design issues that developers need to masterfully build a Webhouse using examples to illustrate key points. Companion Web site features code examples from the book and links to related Web sites.

Design and Implementation

Outlines and Highlights for the Data Warehouse Lifecycle Toolkit by Ralph Kimball, Isbn

Beginning R

Data Warehouse Systems

Mastering Data Warehouse Aggregates

Collaborative Dimensional Modeling, from Whiteboard to Star Schema

This new edition enhances, extends, and clarifies the concepts and examples presented in the first edition. Topics have been restructured to coherently develop the data warehouse architecture.

Market_Desc: · Database/Data Warehouse Developer· Designer or Manager Special Features: · Covers how to design data marts that are well integrated with the overall data warehouse design· Includes CD-ROM with useful checklists· CD-ROM includes useful checklists and an upgraded version of StarTracker software About The Book: In its simplest terms, this book is a step-by-step methodology for designing, developing, and deploying data marts and data warehouses. It shows how dimensional design fits in the overall lifecycle of planning, designing, developing, and deploying data marts and data warehouses. In other words, it covers ALL of the steps a developer needs to go through to guarantee a successful enterprise-wide data warehousing solution. It also covers how to design data marts that are well integrated with the overall data warehouse architecture.

Never HIGHLIGHT a Book Again Virtually all testable terms, concepts, persons, places, and events are included. Cram101 Textbook Outlines gives all of the outlines, highlights, notes for your textbook with optional online practice tests. Only Cram101 Outlines are Textbook Specific. Cram101 is NOT the Textbook. Accompanys: 9780521673761

Cowritten by Ralph Kimball, the world's leading data warehousing authority Delivers real-world solutions for the most time- and labor-intensive portion of data warehousing-data staging, or the extract, transform, load (ETL) process Delineates best practices for extracting data from scattered sources, removing redundant and inaccurate data, transforming the remaining data into correctly formatted data structures, and then loading the end product into the data warehouse Offers proven time-saving ETL techniques, comprehensive guidance on building dimensional structures, and crucial advice on ensuring data quality This book is also available as part of the Kimball's Data Warehouse Toolkit Classics Box Set (ISBN: 9780470479575) with the following 3 books: The Data Warehouse Toolkit, 2nd Edition (9780471200246) The Data Warehouse Lifecycle Toolkit, 2nd Edition (9780470149775) The Data Warehouse ETL Toolkit (9780764567575)

Oracle Data Warehousing and Business Intelligence Solutions

The MicrosoftÂ Data Warehouse Toolkit

Encyclopedia of Database Systems

Clickstream Data Warehousing

THE DATA WAREHOUSE LIFECYCLE TOOLKIT, 2ND ED

ER'99 Workshops on Evolution and Change in Data Management, Reverse Engineering in Information Systems, and the World Wide Web and Conceptual Modeling Paris, France, November 15-18, 1999 Proceedings

The "father of data warehousing" incorporates the latesttechnologies into his blueprint for integrated decision supportsystems Today's corporate IT and data warehouse managers are required tomake a small army of technologies work together to ensure fast andaccurate information for business managers. Bill Inmon created theCorporate Information Factory to solve the needs ofthese managers. Since the First Edition, the design of the factoryhas grown and changed dramatically. This Second Edition, revisedand expanded by 40% with five new chapters, incorporates thesechanges. This step-by-step guide will enable readers to connecttheir legacy systems with the data warehouse and deal with a hostof new and changing technologies, including Web access mechanisms,e-commerce systems, ERP (Enterprise Resource Planning) systems. Thebook also looks closely at exploration and data mining servers foranalyzing customer behavior and departmental data marts forfinance, sales, and marketing.

Updated new edition of Ralph Kimball's groundbreaking book ondimensional modeling for data warehousing and businessintelligence! The first edition of Ralph Kimball's The Data WarehouseToolkit introduced the industry to dimensional modeling,and now his books are considered the most authoritative guides inthis space. This new third edition is a complete library of updateddimensional modeling techniques, the most comprehensive collectionever. It covers new and enhanced star schema dimensional modelingpatterns, adds two new chapters on ETL techniques, includes new andexpanded business matrices for 12 case studies, and more. Authored by Ralph Kimball and Margy Ross, known worldwide aseducators, consultants, and influential thought leaders in datawarehousing and business intelligence Begins with fundamental design recommendations and progressesthrough increasingly complex scenarios Presents unique modeling techniques for business applicationssuch as inventory management, procurement, invoicing, accounting,customer relationship management, big data analytics, and more Draws real-world case studies from a variety of industries,including retail sales, financial services, telecommunications,education, health care, insurance, e-commerce, and more Design dimensional databases that are easy to understand andprovide fast query response with The Data WarehouseToolkit: The Definitive Guide to Dimensional Modeling, 3rdEdition.

Conquer the complexities of this open source statistical language R is fast becoming the de facto standard for statistical computing and analysis in science, business, engineering, and related fields. This book examines this complex language using simple statistical examples, showing how R operates in a user-friendly context. Both students and workers in fields that require extensive statistical analysis will find this book helpful as they learn to use R for simple summary statistics, hypothesis testing, creating graphs, regression, and much more. It covers formula notation, complex statistics, manipulating data and extracting components, and rudimentary programming. R, the open source statistical language increasingly used to handle statistics and produces publication-quality graphs, is notoriously complex This book makes R easier to understand through the use of simple statistical examples, teaching the necessary elements in the context in which R is actually used Covers getting started with R and using it for simple summary statistics, hypothesis testing, and graphs Shows how to use R for formula notation, complex statistics, manipulating data, extracting components, and regression Provides beginning programming instruction for those who want to write their own scripts Beginning R offers anyone who needs to perform statistical analysis the information necessary to use R with confidence.

Big Data Imperatives, focuses on resolving the key questions on everyone's mind: Which data matters? Do you have enough data volume to justify the usage? How you want to process this amount of data? How long do you really need to keep it active for your analysis, marketing, and BI applications? Big data is emerging from the realm of one-off projects to mainstream business adoption; however, the real value of big data is not in the overwhelming size of it, but more in its effective use. This book addresses the following big data characteristics: Very large, distributed aggregations of loosely structured data often incomplete and inaccessible Petabytes/Exabytes of data Millions/billions of people providing/contributing to the context behind the data Flat schema's with few complex interrelationships Involves time-stamped events Made up of incomplete data Includes connections between data elements that must be probabilistically inferred Big Data Imperatives explains 'what big data can do'. It can batch process millions and billions of records both unstructured and structured much faster and cheaper. Big data analytics provide a platform to merge all analysis which enables data analysis to be more accurate, well-rounded, reliable and focused on a specific business capability. Big Data Imperatives describes the complementary nature of traditional data warehouses and big-data analytics platforms and how they feed each other. This book aims to bring the big data and analytics realms together with a greater focus on architectures that leverage the scale and power of big data and the ability to integrate and apply analytics principles to data which earlier was not accessible. This book can also be used as a handbook for practitioners; helping them on methodology,technical architecture, analytics techniques and best practices. At the same time, this book intends to hold the interest of those new to big data and analytics by giving them a deep insight into the realm of big data.

Enterprise Big Data Warehouse, BI Implementations and Analytics

Data Warehousing and Knowledge Discovery

CD-ROM to Accompany The Data Warehouse Lifecycle Toolkit

A Primer for the Data Scientist

9780470149775

Database Design for Mere Mortals

Market_Desc: · Data warehouse Designers· Data warehouse Architects· Data warehouse Developers· Data warehouse Managers Special Features: · The current first edition has sold more than 72,000 copies, generating net revenue of more than $2.5 million· The methods described in this book have been adopted by almost all leading data warehouse vendors· Ralph Kimball and his co-authors are recognized as the driving thought leaders in the data warehousing industry; there is no direct competition· The authors actively promote this methodology in training and consulting worldwide and in their writing in magazines and online About The Book: The book covers best practices from data warehouse project inception through on-going program management. About 30 to 40% of the content in the book is updated and new. This revised tutorial covers major lifecycle topics such as dimensional modeling, tech architecture, ETL, BI etc. It is targeted at both novice and experienced data warehouse professionals.

Data pipelines are the foundation for success in data analytics. Moving data from numerous diverse sources and transforming it to provide context is the difference between having data and actually gaining value from it. This pocket reference defines data pipelines and explains how they work in today's modern data stack. You'll learn common considerations and key decision points when implementing pipelines, such as batch versus streaming data ingestion and build versus buy. This book addresses the most common decisions made by data professionals and discusses foundational concepts that apply to open source frameworks, commercial products, and homegrown solutions. You'll learn: What a data pipeline is and how it works How data is moved and processed on modern data infrastructure, including cloud platforms Common tools and products used by data engineers to build pipelines How pipelines support analytics and reporting needs Considerations for pipeline maintenance, testing, and alerting

"Each chapter is... a practice run for the way we all ought to design our data marts and hence our data warehouses."-Ralph Kimball, from the Foreword. Let the experts show you how to customize data warehouse designs for real business needs in Data Warehouse Design Solutions. To effectively design a data warehouse, you have to understand its many business uses. This guidebook shows you how business managers in different corporate functions actually use data warehouses to make decisions. You'll get a rich set of data warehouse designs that flow from realistic business cases. Two top experts show you how to customize your data warehouse designs for real-life business needs including: * Sales and marketing * Production and inventory management * Budgeting and financial reporting * Quality control * Product delivery and fulfillment * Strategic business analysis such as determining market share, rates of return on investment, and other key analytic ratios. CD-ROM includes All sample data warehouse designs with accompanying preformatted reports in HTML for specific business uses such as marketing, sales, and financial analysis.

Never HIGHLIGHT a Book Again! Virtually all of the testable terms, concepts, persons, places, and events from the textbook are included. Cram101 Just the FACTS101 studyguides give all of the outlines, highlights, notes, and quizzes for your textbook with optional online comprehensive practice tests. Only Cram101 is Textbook Specific. Accompanys: 9780470149775 .

Relational and Dimensional Techniques

THE DATAWAREHOUSE LIFECYCLE TOOLKIT (With CD)

The Data Warehouse Toolkit, 2nd Edition; The Data Warehouse Lifecycle, 2nd Edition; The Data Warehouse ETL Toolkit

With SQL Server 2008 R2 and the Microsoft Business Intelligence Toolset

A Hands-on Guide to Relational Database Design

Building a Data Warehouse for Decision Support

The Data Warehouse Lifecycle ToolkitJohn Wiley & Sons

The data warehousing bible updated for the new millennium Updated and expanded to reflect the many technological advances occurring since the previous edition, this latest edition of the data warehousing "bible" provides a comprehensive introduction to building data marts, operational data stores, the Corporate Information Factory, exploration warehouses, and Web-enabled warehouses. Written by the father of the data warehouse concept, the book also reviews the unique requirements for supporting e-business and explores various ways in which the traditional data warehouse can be integrated with new technologies to provide enhanced customer service, sales, and support-both online and offline-including near-line data storage techniques.

"This book takes the somewhat daunting process of database design and breaks it into completely manageable and understandable components. Mike's approach whilst simple is completely professional, and I can recommend this book to any novice database designer." --Sandra Barker, Lecturer, University of South Australia, Australia "Databases are a critical infrastructure technology for information systems and today's business. Mike Hernandez has written a literate explanation of database technology--a topic that is intricate and often obscure. If you design databases yourself, this book will educate you about pitfalls and show you what to do. If you purchase products that use a database, the book explains the technology so that you can understand what the vendor is doing and assess their products better." --Michael Blaha, consultant and trainer, author of A Manager's Guide to Database Technology "If you told me that Mike Hernandez could improve on the first edition of Database Design for Mere Mortals I wouldn't have believed you, but he did! The second edition is packed with more real-world examples, detailed explanations, and even includes database-design tools on the CD-ROM! This is a must-read for anyone who is even remotely interested in relational database design, from the individual who is called upon occasionally to create a useful tool at work, to the seasoned professional who wants to brush up on the fundamentals. Simply put, if you want to do it right, read this book!" --Matt Greer, Process Control Development, The Dow Chemical Company "Mike's approach to database design is totally common-sense based, yet he's adhered to all the rules of good relational database design. I use Mike's books in my starter database-design class, and I recommend his books to anyone who's interested in learning how to design databases or how to write SQL queries." --Michelle Poolet, President, MVDS, Inc. "Slapping together sophisticated applications with poorly designed data will hurt you just as much now as when Mike wrote his first edition, perhaps even more. Whether you're just getting started developing with data or are a seasoned pro; whether you've read Mike's previous book or this is your first; whether you're happier letting someone else design your data or you love doing it yourself--this is the book for you. Mike's ability to explain these concepts in a way that's not only clear, but fun, continues to amaze me." --From the Foreword by Ken Getz, MCW Technologies, coauthor ASP.NET Developer's JumpStart "The first edition of Mike Hernandez's book Database Design for Mere Mortals was one of the few books that survived the cut when I moved my office to smaller quarters. The second edition expands and improves on the original in so many ways. It is not only a good, clear read, but contains a remarkable quantity of clear, concise thinking on a very complex subject. It's a must for anyone interested in the subject of database design." --Malcolm C. Rubel, Performance Dynamics Associates "Mike's excellent guide to relational database design deserves a second edition. His book is an essential tool for fledgling Microsoft Access and other desktop database developers, as well as for client/server pros. I recommend it highly to all my readers." --Roger Jennings, author of Special Edition Using Access 2002 "There are no silver bullets! Database technology has advanced dramatically, the newest crop of database servers perform operations faster than anyone could have imagined six years ago, but none of these technological advances will help fix a bad database design, or capture data that you forgot to include! Database Design for Mere Mortals(TM), Second Edition, helps you design your database right in the first place!" --Matt Nunn, Product Manager, SQL Server, Microsoft Corporation "When my brother started his professional career as a developer, I gave him Mike's book to help him understand database concepts and make real-world application of database technology. When I need a refresher on the finer points of database design, this is the book I pick up. I do not think that there is a better testimony to the value of a book than that it gets used. For this reason I have wholeheartedly recommended to my peers and students that they utilize this book in their day-to-day development tasks." --Chris Kunicki, Senior Consultant, OfficeZealot.com "Mike has always had an incredible knack for taking the most complex topics, breaking them down, and explaining them so that anyone can 'get it.' He has honed and polished his first very, very good edition and made it even better. If you're just starting out building database applications, this book is a must-read cover to cover. Expert designers will find Mike's approach fresh and enlightening and a source of great material for training others." --John Viescas, President, Viescas Consulting, Inc., author of Running Microsoft Access 2000 and coauthor of SQL Queries for Mere Mortals "Whether you need to learn about relational database design in general, design a relational database, understand relational database terminology, or learn best practices for implementing a relational database, Database Design for Mere Mortals(TM), Second Edition, is an indispensable book that you'll refer to often. With his many years of real-world experience designing relational databases, Michael shows you how to analyze and improve existing databases, implement keys, define table relationships and business rules, and create data views, resulting in data integrity, uniform access to data, and reduced data-entry errors." --Paul Cornell, Site Editor, MSDN Office Developer Center Sound database design can save hours of development time and ensure functionality and reliability. Database Design for Mere Mortals(TM), Second Edition, is a straightforward, platform-independent tutorial on the basic principles of relational database design. It provides a commonsense design methodology for developing databases that work. Database design expert Michael J. Hernandez has expanded his best-selling first edition, maintaining its hands-on approach and accessibility while updating its coverage and including even more examples and illustrations. This edition features a CD-ROM that includes diagrams of sample databases, as well as design guidelines, documentation forms, and examples of the database design process. This book will give you the knowledge and tools you need to create efficient and effective relational databases.

Aimed at helping business and IT managers clearly communicate with each other, this helpful book addresses concerns straight-on and provides practical methods to building a collaborative data warehouse . You'll get clear explanations of the goals and objectives of each stage of the data warehouse lifecycle while learning the roles that both business managers and technicians play at each stage. Discussions of the most critical decision points for success at each phase of the data warehouse lifecycle help you understand ways in which both business and IT management can make decisions that best meet unified objectives.

The Data Warehouse ETL Toolkit

6th International Conference, DaWaK 2004, Zaragoza, Spain, September 1-3, 2004, Proceedings

Agile Data Warehouse Design

The Statistical Programming Language

The Data Warehouse Lifecycle Toolkit

Building the Data Warehouse

Up-to-date, comprehensive coverage of the Oracle database and business intelligence tools Written by a team of Oracle insiders, this authoritative book provides you with the most current coverage of the Oracle data warehousing platform as well as the full suite of business intelligence tools. You'll learn how to leverage Oracle features and how those features can be used to provide solutions to a variety of needs and demands. Plus, you'll get valuable tips and insight based on the authors' real-world experiences and their own implementations. Avoid many common pitfalls while learning best practices for: Leveraging Oracle technologies to design, build, and manage data warehouses Integrating specific database and business intelligence solutions from other vendors Using the new suite of Oracle business intelligence tools to analyze data for marketing, sales, and more Handling typical data warehouse performance challenges Uncovering initiatives by your business community, security business sponsorship, project staffing, and managing risk

Agile Data Warehouse Design is a step-by-step guide for capturing data warehousing/business intelligence (DW/BI) requirements and turning them into high performance dimensional models in the most direct way: by modelstorming (data modeling] brainstorming) with BI stakeholders. This book describes BEAM, an agile approach to dimensional modeling, for improving communication between data warehouse designers, BI stakeholders and the whole DW/BI development team. BEAM provides tools and techniques that will encourage DW/BI designers and developers to move away from their keyboards and entity relationship based tools and model interactively with their colleagues. The result is everyone thinks dimensionally from the outset! Developers understand how to efficiently implement dimensional modeling solutions. Business stakeholders feel ownership of the data warehouse they have created, and can already imagine how they will use it to answer their business questions. Within this book, you will learn: Agile dimensional modeling using Business Event Analysis & Modeling (BEAM) Modelstorming: data modeling that is quicker, more inclusive, more productive, and frankly more fun! Telling dimensional data stories using the 7Ws (who, what, when, where, how many, why and how) Modeling by example not abstraction; using data story themes, not crow's feet, to describe detail Storyboarding the data warehouse to discover conformed dimensions and plan iterative development Visual modeling: sketching timelines, charts and grids to model complex process measurement - simply Agile design documentation: enhancing star schemas with BEAM dimensional shorthand notation Solving difficult DW/BI performance and usability problems with proven dimensional design patterns LawrenceCorr is a data warehouse designer and educator. As Principal of DecisionOne Consulting, he helps clients to review and simplify their data warehouse designs, and advises vendors on visual data modeling techniques. He regularly teaches agile dimensional modeling courses worldwide and has taught dimensional DW/BI skills to thousands of students. Jim Stagnitto is a data warehouse and master data management architect specializing in the healthcare, financial services, and information service industries. He is the founder of the data warehousing and data mining consulting firm Llumino.

Three books by the bestselling authors on Data Warehousing! The most authoritative guides from the inventor of the technique all for a value price. The Data Warehouse Toolkit, 3rd Edition (9781118530801) Ralph Kimball invented a data warehousing technique called "dimensional modeling" and popularized it in his first Wiley book, The Data Warehouse Toolkit. Since this book was first published in 1996, dimensional modeling has become the most widely accepted technique for data warehouse design. Over the past 10 years, Kimball has improved on his earlier techniques and created many new ones. In this 3rd edition, he will provide a comprehensive collection of all of these techniques, from basic to advanced. The Data Warehouse Lifecycle Toolkit, 2nd Edition (9780470149775) Complete coverage of best practices from data warehouse project inception through on-going program management. Updates industry best practices to be in sync with current recommendations of Kimball Group. Streamlines the lifecycle methodology to be more efficient and user-friendly The Data Warehouse ETL Toolkit (9780764567575) shows data warehouse developers how to effectively manage the ETL (Extract, Transform, Load) phase of the data warehouse development lifecycle. The authors show developers the best methods for extracting data from scattered sources throughout the enterprise, removing obsolete, redundant, and innaccurate data, transforming the remaining data into correctly formatted data structures, and then physically loading them into the data warehouse. This book provides complete coverage of proven, time-saving ETL techniques. It begins with a quick overview of ETL fundamentals and the role of the ETL development team. It then quickly moves into an overview of the ETL data structures, both relational and dimensional. The authors show how to build useful dimensional stuctures, providing practical examples of beginning through advanced techniques.

A thorough update to the industry standard for designing, developing, and deploying data warehouse and business intelligence systems The world of data warehousing has changed remarkably since the first edition of The Data Warehouse Lifecycle Toolkit was published in 1998. In that time, the data warehouse industry has reached full maturity and acceptance, hardware and software have made staggering advances, and the techniques promoted in the premiere edition of this book have been adopted by nearly all data warehouse vendors and practitioners. In addition, the term "business intelligence" emerged to reflect the mission of the data warehouse: wrangling the data out of source systems, cleaning it, and delivering it to add value to the business. Ralph Kimball and his colleagues have refined the original set of Lifecycle methods and techniques based on their consulting and training experience. The authors understand first-hand that a data warehousing/business intelligence (DW/BI) system needs to change as fast as its surrounding organization evolves. To that end, they walk you through the detailed steps of designing, developing, and deploying a DW/BI system. You'll learn to create adaptable systems that deliver data and analyses to business users so they can make better business decisions.

The Microsoft Data Warehouse Toolkit

With Examples in SQL Server

Corporate Information Factory

Three Volume Set of Ralph Kimball's Toolkit Books

The Definitive Guide to Dimensional Modeling

T-SQL in One Hour a Day, Sams Teach Yourself

With this textbook, Vaisman and Zimányi deliver excellent coverage of data warehousing and business intelligence technologies ranging from the most basic principles to recent findings and applications. To this end, their work is structured into three parts. Part I describes “Fundamental Concepts” including conceptual and logical data warehouse design, as well as querying using MDX, DAX and SQL/OLAP. This part also covers data analytics using Power BI and Analysis Services. Part II details “Implementation and Deployment,” including physical design, ETL and data warehouse design methodologies. Part III covers “Advanced Topics” and it is almost completely new in this second edition. This part includes chapters with an in-depth coverage of temporal, spatial, and mobility data warehousing. Graph data warehouses are also covered in detail using Neo4j. The last chapter extensively studies big data management and the usage of Hadoop, Spark, distributed, in-memory, columnar, NoSQL and NewSQL database systems, and data lakes in the context of analytical data processing. As a key characteristic of the book, most of the topics are presented and illustrated using application tools. Specifically, a case study based on the well-known Northwind database illustrates how the concepts presented in the book can be implemented using Microsoft Analysis Services and Power BI. All chapters have been revised and updated to the latest versions of the software tools used. KPIs and Dashboards are now also developed using DAX and Power BI, and the chapter on ETL has been expanded with the implementation of ETL processes in PostgreSQL. Review questions and exercises complement each chapter to support comprehensive student learning. Supplemental material to assist instructors using this book as a course text is available online and includes electronic versions of the figures, solutions to all exercises, and a set of slides accompanying each chapter. Overall, students, practitioners and researchers alike will find this book the most comprehensive reference work on data warehouses, with key topics described in a clear and educational style. “I can only invite you to dive into the contents of the book, feeling certain that once you have completed its reading (or maybe, targeted parts of it), you will join me in expressing our gratitude to Alejandro and Esteban, for providing such a comprehensive textbook for the field of data warehousing in the first place, and for keeping it up to date with the recent developments, in this current second edition.” From the foreword by Panos Vassiliadis, University of Ioannina, Greece.

Over the past 5 years, the concept of big data has matured, data science has grown exponentially, and data architecture has become a standard part of organizational decision-making. Throughout all this change, the basic principles that shape the architecture of data have remained the same. There remains a need for people to take a look at the "bigger picture" and to understand where their data fit into the grand scheme of things. Data Architecture: A Primer for the Data Scientist, Second Edition addresses the larger architectural picture of how big data fits within the existing information infrastructure or data warehousing systems. This is an essential topic not only for data scientists, analysts, and managers but also for researchers and engineers who increasingly need to deal with large and complex sets of data. Until data are gathered and can be placed into an existing framework or architecture, they cannot be used to their full potential. Drawing upon years of practical experience and using numerous examples and case studies from across various industries, the authors seek to explain this larger picture into which big data fits, giving data scientists the necessary context for how pieces of the puzzle should fit together. New case studies include expanded coverage of textual management and analytics New chapters on visualization and big data Discussion of new visualizations of the end-state architecture

Focus your efforts on the best opportunities --

An evidence-based organizational framework for exceptional analytics team results The Analytics Lifecycle Toolkit provides managers with a practical manual for integrating data management and analytic technologies into their organization. Author Gregory Nelson has encountered hundreds of unique perspectives on analytics optimization from across industries; over the years, successful strategies have proven to share certain practices, skillsets, expertise, and structural traits. In this book, he details the concepts, people and processes that contribute to exemplary results, and shares an organizational framework for analytics team functions and roles. By merging analytic culture with data and technology strategies, this framework creates understanding for analytics leaders and a toolbox for practitioners. Focused on team effectiveness and the design thinking surrounding product creation, the framework is illustrated by real-world case studies to show how effective analytics team leadership works on the ground. Tools and templates include best practices for process improvement, workforce enablement, and leadership support, while guidance includes both conceptual discussion of the analytics life cycle and detailed process descriptions. Readers will be equipped to: Master fundamental concepts and practices of the analytics life cycle Understand the knowledge domains and best practices for each stage Delve into the details of analytical team processes and process optimization Utilize a robust toolkit designed to support analytic team effectiveness The analytics life cycle includes a diverse set of considerations involving the people, processes, culture, data, and technology, and managers needing stellar analytics performance must understand their unique role in the process of winnowing the big picture down to meaningful action. The Analytics Lifecycle Toolkit provides expert perspective and much-needed insight to managers, while providing practitioners with a new set of tools for optimizing results.

Advances in Conceptual Modeling

The Complete Guide to Dimensional Modeling

A Practical Guide for an Effective Analytics Capability

Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data

Building a Data Warehouse

Big Data Imperatives

Cowritten by Ralph Kimball, the world's leading data warehousing authority, whose previous books have sold more than 150,000 copies Delivers real-world solutions for the most time- and labor-intensive portion of data warehousing-data staging, or the extract, transform, load (ETL) process Delineates best practices for extracting data from scattered sources, removing redundant and inaccurate data, transforming the remaining data into correctly formatted data structures, and then loading the end product into the data warehouse Offers proven time-saving ETL techniques, comprehensive guidance on building dimensional structures, and crucial advice on ensuring data quality

Here is the ideal field guide for data warehousing implementation. This book first teaches you how to build a data warehouse, including defining the architecture, understanding the methodology, gathering the requirements, designing the data models, and creating the databases. Coverage then explains how to populate the data warehouse and explores how to present data to users using reports and multidimensional databases and how to use the data in the data warehouse for business intelligence, customer relationship management, and other purposes. It also details testing and how to administer data warehouse operation.

Master T-SQL database design, development, and administration the easy way–hands-on! In just one hour a day, you’ll build all the skills you need to create effective database applications with T-SQL and SQL Server. With this complete tutorial, you’ll quickly master the basics and then move on to more advanced features and concepts: Learn the fundamentals of T-SQL from the ground up, one step at a time Succeed with the newest versions of T-SQL, SQL Server, and SQL Server Management Studio Use T-SQL effectively as both an application developer and DBA Master powerful stored procedures, triggers, transactions, and user-defined functions (UDFs) Systematically optimize and secure your SQL Server databases Learn on your own time, at your own pace No previous T-SQL or database programming experience required Learn how to design efficient, reliable SQL Server databases Define efficient tables, table relationships, fields, and constraints Make the most of T-SQL’s SELECT and UPDATE statements Work effectively with simple and complex views and joins Master stored procedure techniques every developer should know Build and use powerful User-Defined Functions (UDFs) Secure databases with authentication, roles, permissions, and principals Configure, maintain, and tune SQL Server for maximum reliability, performance, and value Back up, restore, and audit databases Optimize databases with the SQL Server Profiler, System Monitor, and Index Tuning Wizard Leverage valuable insight and time saving techniques from a world renowned database expert Register your book at informit.com/register for access to source code, example files, updates, and corrections as they become available.

Data Architecture: A Primer for the Data Scientist

The Data Warehouse Toolkit

3 Volume Set

Solutions for Star Schema Performance

Data Warehouse Performance