The Data Warehouse 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 Toolkit (9780764567757) shows data 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 inaccurate 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 structures, providing practical examples of beginning through advanced techniques.

The Microsoft Data Warehouse Toolkit offers proven time-saving ETL techniques, comprehensive guidance on building dimensional structures, and crucial advice on ensuring data quality. It offers proven time-saving ETL techniques, comprehensive guidance on building dimensional structures, and crucial advice on ensuring data quality.
ComputerWorld named Bill one of the ten most influential people in the history of the computer profession. Bill’s latest adventure is the building of technology known as textual disambiguation. About architecture and when to apply contextualization and textual disambiguation. About Bill Inmon, the “father of the data warehouse,” has written 60 books published in nine languages. document preprocessing. Appreciate the power of taxonomies and the necessity of textual ETL. Know how the textual warehouse architecture differs from the conventional data warehouse analysis), to make better business decisions. Learn the important role of documents and text within your organization, the difference between identifying and qualifying text, and when you need Lake? Data Lake Architecture Key Data Lake Concepts Maturity stages of Data Lake Chapter 14: Data Lake Vs Data Warehouse What is Data Warehouse? What is Data Lake? Key Difference between Schemas? What is a Star Schema? What is a Snowflake Schema? Difference between Star Schema and Snowflake Chapter 11: Data Mart What is Data Mart? Type of Data Mart Steps in Implementing can be completed in a day. With this e-book, you will be enough knowledge to contribute and participate in a Data warehouse implementation project. The book covers upcoming and promising project planning, requirements gathering, dimensional modeling, ETL, and business intelligence and analytics. This easily referenced collection is nothing less than vital if you are involved with data lifecycle methodology developed by the Kimball Group. These practical, hands-on articles are fully updated to reflect current practices and terminology and cover the complete lifecycle—including project planning, requirements gathering, dimensional modeling, ETL, and business intelligence and analytics. This easily referenced collection is nothing less than vital if you are involved with data warehousing. Writes 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

Three Volume Set of Ralph Kimball’s Toolkit Books-Ralph Kimball 2000-06-01 Ralph Kimball’s three data warehouse books, The Data Warehouse Toolkit, The Data Warehouse Lifecycle Toolkit, and The Data Warehouse Bookshelf, 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 discusses case studies of existing data warehouse projects for specific types of data such as retail, manufacturing, banking, insurance, subscriptions 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 Warehouse Bookshelf is a groundbreaking guide that 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.

Mastering Data Warehouse Aggregates-Christopher Adamson 2012-06-27 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.

Data Warehouse Solutions-Christopher Adamson 1996-07-13 “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 details, examples, and tips for designing a data warehouse architecture and then building the components that make up a data warehouse. Examine the key issues surrounding data warehousing and the many functional components of a data warehouse. The book offers a thorough grounding in the important role of documents and text within your organization, the difference between identifying and qualifying text, and when you need Lake? Data Lake Architecture Key Data Lake Concepts Maturity stages of Data Lake Chapter 14: Data Lake Vs Data Warehouse What is Data Warehouse? What is Data Lake? Type of Data Lake Steps in Implementing can be completed in a day. With this e-book, you will be enough knowledge to contribute and participate in a Data warehouse implementation project. The book covers upcoming and promising project planning, requirements gathering, dimensional modeling, ETL, and business intelligence and analytics. This easily referenced collection is nothing less than vital if you are involved with data warehousing. Writes 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

Learn Data Warehousing in 1 Day-Krishna Rungta 2018-02-15 Unlike popular tool-based popular belief, Data Warehouse is not a single tool but a collection of software tools. A data warehouse will collect data from diverse sources into a single database. Using Business Intelligence tools, meaningful insights are drawn from this data. The best thing about “Learn Data Warehousing in 1 Day” is that it is small and can be read in one day. With this book, you’ll be well on your way in building a knowledge warehouse implementation project. The book covers the following topics: dimensional modeling, ETL, and business intelligence and analytics. This easily referenced collection is nothing less than vital if you are involved with data warehousing or business intelligence in any capacity.

The Kimball Group Reader-Ralph Kimball 2010-03-11 An unparalleled collection of recommended guidelines for data warehousing and business intelligence pioneered by Ralph Kimball and his team of colleagues from the Kimball Group. Recognized and respected throughout the world as the most influential leaders in the data warehousing industry, Ralph Kimball and the Kimball Group have written articles covering more than 250 topics that define the field of data warehousing. For the first time, the Kimball Group’s incomparable advice, design tips, and best practices have been gathered in this remarkable collection of articles, which spans a decade of data warehousing innovation. Each group of articles is introduced with original commentaries that explain their role in the overall lifecycle methodology developed by the Kimball Group. These practical, hands-on articles are fully updated to reflect current practices and terminology and cover the complete lifecycle—including project planning, requirements gathering, dimensional modeling, ETL, and business intelligence and analytics. This easily referenced collection is nothing less than vital if you are involved with data warehousing or business intelligence in any capacity.

The Textual Warehouse-Bill Inmon 2021-07-12 Build a Textual Warehouse to help your organization understand and analyze documents through text analytics (both sentiment and non-sentiment analysis), to make better business decisions. Learn the important role of documents and text within your organization, the difference between identifying and qualifying text, and when you need Lake? Data Lake Architecture Key Data Lake Concepts Maturity stages of Data Lake Chapter 14: Data Lake Vs Data Warehouse What is Data Warehouse? What is Data Lake? Type of Data Lake Steps in Implementing can be completed in a day. With this e-book, you will be enough knowledge to contribute and participate in a Data warehouse implementation project. The book covers upcoming and promising project planning, requirements gathering, dimensional modeling, ETL, and business intelligence and analytics. This easily referenced collection is nothing less than vital if you are involved with data warehousing or business intelligence in any capacity.

The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming and Delivering Ralph Kimball 2010-03-11 The definitive guide to building a data warehouse. Kimball discusses case studies of existing data warehouse projects for specific types of data such as retail, manufacturing, banking, insurance, subscriptions 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 Warehouse Bookshelf is a groundbreaking guide that 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.

Three Volume Set of Ralph Kimball’s Toolkit Books-Ralph Kimball 2000-06-01 Ralph Kimball’s three data warehouse books, The Data Warehouse Toolkit, The Data Warehouse Lifecycle Toolkit, and The Data Warehouse Bookshelf, 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 discusses case studies of existing data warehouse projects for specific types of data such as retail, manufacturing, banking, insurance, subscriptions 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 Warehouse Bookshelf is a groundbreaking guide that 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.

The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming and Delivering Ralph Kimball 2010-03-11 The definitive guide to building a data warehouse. Kimball discusses case studies of existing data warehouse projects for specific types of data such as retail, manufacturing, banking, insurance, subscriptions 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 Warehouse Bookshelf is a groundbreaking guide that 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.

The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming and Delivering Ralph Kimball 2010-03-11 The definitive guide to building a data warehouse. Kimball discusses case studies of existing data warehouse projects for specific types of data such as retail, manufacturing, banking, insurance, subscriptions 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 Warehouse Bookshelf is a groundbreaking guide that 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.
Readers will learn the techniques, technologies, and best practices for managing the passage of data between computer systems and integrating disparate data together in an enterprise environment. The average enterprise's computing environment is comprised of hundreds to thousands computer systems that have been built, purchased, and acquired over time. The data from these various systems needs to be integrated for reporting and analysis, shared for business transaction processing, and converted from one format to another when old systems are replaced and new systems are acquired. The management of the "data in motion" in organizations is rapidly becoming one of the biggest concerns for business and IT management. Data warehousing and conversion, real-time data integration, and cloud and "big data" applications are just a few of the challenges facing organizations and businesses today. Managing data in motion tackles these and other topics in a style easily understood by business and IT managers as well as programmers and architects. Presents a summary of the different technologies and techniques for moving data between computer systems including the emerging solutions for unstructured as well as structured data types. \textbf{Explains, in non-technical terms, the architecture and components required to perform data integration.} Describes how to reduce the complexity of managing system interfaces and enable a scalable data architecture that can handle the dimensions of "Big Data".

**Data Warehousing**

Paul Westerman 2001

What is data warehousing? 

- Project planning 
- Business exploration 
- Business case study and ROI analysis 
- Organizational integration 
- Technology 
- Database maintenance 
- Technical construction of the Wal-Mart data warehouse 
- Post-implementation of the Wal-Mart data warehouse 
- Store operations sample analyses 
- Merchandising sample analyses.

**SQL & PL/SQL FOR ORACLE 11G BLACK BOOK (With CD)**

P.S.Deshpande 2011-07-01

Special Features: Dive into the world of SQL and PL/SQL programming in Oracle 10g: Grasp all theoretical and practical concepts in programming. Understand the advanced features of SQL, like Analytical queries, multitable inserts, joins, Scalar subqueries, key preserving joins, etc. Understand the concept of Server model in software. Place OCP exams and interviews (for software professionals). Absorb advanced features in PL/SQL, like Control structures, Collections, LR0 datatype, Object types, Error handling, Triggers, Object views, Rollback segments, UNDO/Redo space, Flashback queries, Sequence and Wrapper, Expression filter, in addition to Security and Audit. About The Book: This book, SQL & PL/SQL Black Book, brings to you the complete, latest account of SQL and PL/SQL for Oracle. It gifts you two things-first, it upgrades your knowledge on the subject; and second, the innumerable, exciting examples help you build a better skill in SQL and PL/SQL. The highlight of the book is a clear yet incisive discussion of the basic features of Oracle 9i and the latest advanced features that come with Oracle 10g. The in-depth treatment of the subject in a lucid language will give you an enjoyable learning experience as you have always had with our previous books on Oracle. This comprehensive reference will help you learn the entire spectrum of SQL- from Views to Caeve and Roll Up; from case expressions to Variances and Correlation; PL/SQL from Object types; from Error handling to Triggers; from Object views to Rollback segments; from UNDO/Redo space to Expression filter, and many more. All this, as the book zips through the material and does not blather on or repeat points made earlier-no doubt, every aspect is worth the price of the entire book.

**Fundamentals of Data Warehouses**

Matthias Jarke 2013-03-09

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers modern data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

**Data Quality**

Jack E. Olson 2003-01-09

Data Quality: The Accuracy Dimension is about assessing the quality of corporate data and improving its accuracy using the data profiling method. Corporate data is increasingly important as companies continue to find new ways to use it. Likewise, improving the accuracy of data in information systems is fast becoming a major goal as companies realize how much it affects their bottom line. Data profiling is a new technology that supports and enhances the accuracy of databases throughout major IT shops. Jack Olson explains data profiling and shows how it fits into the larger picture of data quality. * Provides an accessible, enjoyable introduction to the subject of data accuracy, peppered with real-world anecdotes. * Provides a framework for data profiling with a discussion of analytical tools appropriate for assessing data accuracy. * Is written by one of the original developers of data profiling technology. * Is a must-read for any data management staff, IT management staff, and CIOs of companies with data assets.

**Big Data Imperatives**

Soumendra Mohanty 2013-08-23

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 expense? How much 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 are used, covering trillions of records that require processing to understand different aspects of the business. Big data analytics provide a platform to merge all large data sources that 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 both 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.

**Building a Scalable Data Warehouse with Data Vault 2.0**

Dan Linstedt 2015-09-15

The Data Vault was invented by Dan Linstedt at the U.S. Department of Defense, and the standard has been successfully applied to data warehousing projects at organizations of different sizes, from small to large-size corporations. Due to its simplified design, which is adapted from nature, the Data Vault 2.0 standard helps prevent typical data warehousing failures. "Building a Scalable Data Warehouse" covers everything one needs to know to create a scalable data warehouse end to end, including a presentation of the Data Vault modeling technique, which provides the foundations to create a technical data warehouse layer. The book discusses how to build the data warehouse incrementally using the agile Data Vault 2.0 methodology. In addition, readers will learn how to create the input layer (the stage layer) and the presentation layer (data mart) of the Data Vault 2.0 architecture including implementation best practices. Drawing upon years of practical experience and using numerous examples and an easy to understand framework, Dan Linstedt and Michael Olschimke discuss how to load each layer using SQL Server Integration Services (SSIS), including automation of the Data Vault loading processes. Important data warehouse technologies and practices. Data Quality Services (DQS) and Master Data Services (MDS) in the context of the Data Vault architecture. Provides a complete introduction to data warehousing, applications, and the business context so readers can get-up and running fast Explains theoretical concepts and provides hands-on instruction on how to build and implement a data warehouse. Demystifies data vault modeling with beginning, intermediate, and advanced techniques Discusses the advantages of the data vault approach over other techniques, also including the latest updates to Data Vault 2.0 and multiple improvements to Data Vault 1.0.