

The Data Warehouse Mentor Practical Data Warehouse And Business Intelligence Insights

Yeah, reviewing a ebook **The Data Warehouse Mentor Practical Data Warehouse And Business Intelligence Insights** could be credited with your near links listings. This is just one of the solutions for you to be successful. As understood, skill does not suggest that you have astonishing points.

Comprehending as capably as contract even more than extra will provide each success. next-door to, the proclamation as without difficulty as keenness of this The Data Warehouse Mentor Practical Data Warehouse And Business Intelligence Insights can be taken as without difficulty as picked to act.

Microsoft SQL Server 2008 Integration Services - Erik Veerman 2009-10-16

An authoritative guide to designing effective solutions for datacleansing, ETL, and file management with SQL Server 2008Integration Services SQL Server Integration Services (SSIS) is the leading tool inthe data warehouse industry, used for performing extraction,transformation, and load operations. After an overview of SSIS architecture, the authors walk you aseries of real-world problems and show various techniques forhandling them. Shows you how to design SSIS solutions for data cleansing, ETLand file management Demonstrates how to integrate data from a variety of datasources, Shows how to monitor SSIS performance, Demonstrates how to avoid common pitfalls involved with SSISdeployment Explains how to ensure performance of the deployed solution andeffectively handle unexpected system failures and outages The companion Web site provides sample code and databasescripts that readers can directly implement This book shows you how to design, build, deploy, and managesolutions to real-world problems that SSIS administrators anddevelopers face day-to-day.

Practical Data Mining - Jr. Hancock 2019-09-23

Used by corporations, industry, and government to inform and fuel everything from focused advertising to homeland security, data mining can be a very useful tool across a wide range of applications. Unfortunately, most books on the subject are designed for the computer scientist and statistical illuminati and leave the reader largely adrift in technical waters. Revealing the lessons known to the seasoned expert, yet rarely written down for the uninitiated, Practical Data Mining explains the ins-and-outs of the detection, characterization, and exploitation of actionable patterns in data. This working field manual outlines the what, when, why, and how of data mining and offers an easy-to-follow, six-step spiral process. Catering to IT consultants, professional data analysts, and sophisticated data owners, this systematic, yet informal treatment will help readers answer questions, such as: What process model should I use to plan and execute a data mining project? How is a quantitative business case developed and assessed? What are the skills needed for different data mining projects? How do I track and evaluate data mining projects? How do I choose the best data mining techniques? Helping you avoid common mistakes, the book describes specific genres of data mining practice. Most chapters contain one or more case studies with detailed projects descriptions, methods used, challenges encountered, and results obtained. The book includes working checklists for each phase of the data mining process. Your passport to successful technical and planning discussions with management, senior scientists, and customers, these checklists lay out the right questions to ask and the right points to make from an insider's point of view. Visit the book's webpage

Computerworld - 2007-01-15

For more than 40 years, Computerworld has been the leading source of technology news and information for IT influencers worldwide. Computerworld's award-winning Web site (Computerworld.com), twice-monthly publication, focused conference series and custom research form the hub of the world's largest global IT media network.

Data Warehousing and Analytics - David Taniar 2022-03-06

This textbook covers all central activities of data warehousing and analytics, including transformation, preparation, aggregation, integration, and analysis. It discusses the full spectrum of the journey of data from operational/transactional databases, to data warehouses and data analytics; as well as the role that data warehousing plays in the data processing lifecycle. It also explains in detail how data warehouses may

be used by data engines, such as BI tools and analytics algorithms to produce reports, dashboards, patterns, and other useful information and knowledge. The book is divided into six parts: "Part I - Star Schema" describes the foundation of data warehouse design. "Part II - Snowflake and Bridge Tables" then expands the concept of a simple star schema by introducing the concept of hierarchy, bridge tables, as well as the use of bridge tables in temporal data warehousing. "Part III - Advanced Dimensions" elaborates various dimension models, namely determinant dimensions, junk dimensions, dimension keys, and one-attribute dimensions, which all enrich the semantics of the star schema. "Part IV - Multi-Fact and Multi-Input" introduces multi-fact star schemas, where the star schema has multi-fact entities. A multi-fact can also be created by slicing one fact into multi-facts, which is discussed next. Eventually the creation of a star schema is introduced where an operational database is used as input to the transformation process and consists of multiple operational databases. "Part V - Data Warehousing Granularity and Evolution" first introduces the concept of aggregation levels in a star schema constellation. It then focuses on the lowest-level star schema, including how to design a star schema and why it is needed in data warehousing, before moving on to methods for adding and removing dimensions. The remaining two chapters present more advanced concepts in data warehousing granularity and introduce the concept of active data warehousing. "Part VI - OLAP, Business Intelligence, and Data Analytics" thoroughly explains OLAP - online analytical processing, and describes two important activities in the data warehousing process, namely pre-data warehousing and post-data warehousing. The final chapter focuses on data analytics, which consists of a suite of methods for data analysis suitable for data warehousing. This textbook approaches data warehousing from the case study angle. Each chapter presents one or more case studies to thoroughly explain the concepts and has different levels of difficulty, hence learning is incremental. In addition, every chapter has also a section on further readings which give pointers and references to research papers related to the chapter. All these features make the book ideally suited for either introductory courses on data warehousing and data analytics, or even for self-studies by professionals. The book is accompanied by a web page that includes all the used datasets and codes as well as slides and solutions to exercises.

Learning Informatica PowerCenter 10.x - Rahul Malewar 2017-08-10

Harness the power and simplicity of Informatica PowerCenter 10.x to build and manage efficient data management solutions About This Book Master PowerCenter 10.x components to create, execute, monitor, and schedule ETL processes with a practical approach. An ideal guide to building the necessary skills and competencies to become an expert Informatica PowerCenter developer. A comprehensive guide to fetching/transforming and loading huge volumes of data in a very effective way, with reduced resource consumption Who This Book Is For If you wish to deploy Informatica in enterprise environments and build a career in data warehousing, then this book is for you. Whether you are a software developer/analytic professional and are new to Informatica or an experienced user, you will learn all the features of Informatica 10.x. A basic knowledge of programming and data warehouse concepts is essential. What You Will Learn Install or upgrade the components of the Informatica PowerCenter tool Work on various aspects of administrative skills and on the various developer Informatica PowerCenter screens such as Designer, Workflow Manager, Workflow Monitor, and Repository Manager. Get practical hands-on experience of various sections of Informatica PowerCenter, such as navigator, toolbar, workspace, control panel, and so on Leverage basic and advanced utilities, such as the debugger, target load plan, and incremental

aggregation to process data Implement data warehousing concepts such as schemas and SCDs using Informatica Migrate various components, such as sources and targets, to another region using the Designer and Repository Manager screens Enhance code performance using tips such as pushdown optimization and partitioning In Detail Informatica PowerCenter is an industry-leading ETL tool, known for its accelerated data extraction, transformation, and data management strategies. This book will be your quick guide to exploring Informatica PowerCenter's powerful features such as working on sources, targets, transformations, performance optimization, scheduling, deploying for processing, and managing your data at speed. First, you'll learn how to install and configure tools. You will learn to implement various data warehouse and ETL concepts, and use PowerCenter 10.x components to build mappings, tasks, workflows, and so on. You will come across features such as transformations, SCD, XML processing, partitioning, constraint-based loading, Incremental aggregation, and many more. Moreover, you'll also learn to deliver powerful visualizations for data profiling using the advanced monitoring dashboard functionality offered by the new version. Using data transformation technique, performance tuning, and the many new advanced features, this book will help you understand and process data for training or production purposes. The step-by-step approach and adoption of real-time scenarios will guide you through effectively accessing all core functionalities offered by Informatica PowerCenter version 10.x. Style and approach You'll get hand-on with sources, targets, transformations, performance optimization, scheduling, deploying for processing, and managing your data, and learn everything you need to become a proficient Informatica PowerCenter developer.

Mentor Coaching: A Practical Guide - Clare Norman 2020-03-26

This practical guide argues that both mentor coaching and supervision be mandated by the professional coaching bodies as part of coaches' continuous professional development. Mentor coaching is not just for those coaches seeking a credential: it is for lifelong professional development for every coach, at every level of the profession. You are the best coaching tool there is. Tools need to be oiled, sharpened, repaired and protected to keep them in tip top condition. That's what mentor coaching and supervision do - they keep coaches sharp and fit for purpose. The reader will learn how to develop as a coach using mentor coaching, as well as how to develop as a mentor coach, to support other coaches to develop. Clare Norman explains what mentor coaching is, why it is so important and the competencies for mentor coaching. How coaches show up in the room is more important than how much we know about the theory behind coaching. Clare Norman's concise book is important reading for all practising coaches, as well as coaches and mentor coaches in training.

The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights -

Robert Laberge 2011-06-05

Develop a custom, agile data warehousing and business intelligence architecture Empower your users and drive better decision making across your enterprise with detailed instructions and best practices from an expert developer and trainer. The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights shows how to plan, design, construct, and administer an integrated end-to-end DW/BI solution. Learn how to choose appropriate components, build an enterprise data model, configure data marts and data warehouses, establish data flow, and mitigate risk. Change management, data governance, and security are also covered in this comprehensive guide. Understand the components of BI and data warehouse systems Establish project goals and implement an effective deployment plan Build accurate logical and physical enterprise data models Gain insight into your company's transactions with data mining Input, cleanse, and normalize data using ETL (Extract, Transform, and Load) techniques Use structured input files to define data requirements Employ top-down, bottom-up, and hybrid design methodologies Handle security and optimize performance using data governance tools Robert Laberge is the founder of several Internet ventures and a principle consultant for the IBM Industry Models and Assets Lab, which has a focus on data warehousing and business intelligence solutions.

Critical Mentoring - Torie Weiston-Serdan 2017-02-28

This book introduces the concept of critical mentoring, presenting its theoretical and empirical foundations, and providing telling examples of what it looks like in practice, and what it can achieve. At this juncture when the demographics of our schools and colleges are rapidly changing, critical mentoring provides

mentors with a new and essential transformational practice that challenges deficit-based notions of protégés, questions their forced adaptation to dominant ideology, counters the marginalization and minoritization of young people of color, and endows them with voice, power and choice to achieve in society while validating their culture and values. Critical mentoring places youth at the center of the process, challenging norms of adult and institutional authority and notions of saviorism to create collaborative partnerships with youth and communities that recognize there are multiple sources of expertise and knowledge. Torie Weiston-Serdan outlines the underlying foundations of critical race theory, cultural competence and intersectionality, describes how collaborative mentoring works in practice in terms of dispositions and structures, and addresses the implications of rethinking about the purposes and delivery of mentoring services, both for mentors themselves and the organizations for which they work. Each chapter ends with a set of salient questions to ask and key actions to take. These are meant to move the reader from thought to action and provide a basis for discussion. This book offers strategies that are immediately applicable and will create a process that is participatory, emancipatory and transformative.

Relational Database Design and Implementation - Jan L. Harrington 2016-04-15

Relational Database Design and Implementation: Clearly Explained, Fourth Edition, provides the conceptual and practical information necessary to develop a database design and management scheme that ensures data accuracy and user satisfaction while optimizing performance. Database systems underlie the large majority of business information systems. Most of those in use today are based on the relational data model, a way of representing data and data relationships using only two-dimensional tables. This book covers relational database theory as well as providing a solid introduction to SQL, the international standard for the relational database data manipulation language. The book begins by reviewing basic concepts of databases and database design, then turns to creating, populating, and retrieving data using SQL. Topics such as the relational data model, normalization, data entities, and Codd's Rules (and why they are important) are covered clearly and concisely. In addition, the book looks at the impact of big data on relational databases and the option of using NoSQL databases for that purpose. Features updated and expanded coverage of SQL and new material on big data, cloud computing, and object-relational databases Presents design approaches that ensure data accuracy and consistency and help boost performance Includes three case studies, each illustrating a different database design challenge Reviews the basic concepts of databases and database design, then turns to creating, populating, and retrieving data using SQL

Computerworld - 2003-04-14

For more than 40 years, Computerworld has been the leading source of technology news and information for IT influencers worldwide. Computerworld's award-winning Web site (Computerworld.com), twice-monthly publication, focused conference series and custom research form the hub of the world's largest global IT media network.

Data Architecture - Charles Tupper 2011-05-09

Data Architecture: From Zen to Reality explains the principles underlying data architecture, how data evolves with organizations, and the challenges organizations face in structuring and managing their data. Using a holistic approach to the field of data architecture, the book describes proven methods and technologies to solve the complex issues dealing with data. It covers the various applied areas of data, including data modelling and data model management, data quality, data governance, enterprise information management, database design, data warehousing, and warehouse design. This text is a core resource for anyone customizing or aligning data management systems, taking the Zen-like idea of data architecture to an attainable reality. The book presents fundamental concepts of enterprise architecture with definitions and real-world applications and scenarios. It teaches data managers and planners about the challenges of building a data architecture roadmap, structuring the right team, and building a long term set of solutions. It includes the detail needed to illustrate how the fundamental principles are used in current business practice. The book is divided into five sections, one of which addresses the software-application development process, defining tools, techniques, and methods that ensure repeatable results. Data Architecture is intended for people in business management involved with corporate data issues and information technology decisions, ranging from data architects to IT consultants, IT auditors, and data

administrators. It is also an ideal reference tool for those in a higher-level education process involved in data or information technology management. Presents fundamental concepts of enterprise architecture with definitions and real-world applications and scenarios Teaches data managers and planners about the challenges of building a data architecture roadmap, structuring the right team, and building a long term set of solutions Includes the detail needed to illustrate how the fundamental principles are used in current business practice

Business Intelligence Roadmap - Larissa T. Moss 2003-02-25

"If you are looking for a complete treatment of business intelligence, then go no further than this book. Larissa T. Moss and Shaku Atre have covered all the bases in a cohesive and logical order, making it easy for the reader to follow their line of thought. From early design to ETL to physical database design, the book ties together all the components of business intelligence." --Bill Inmon, Inmon Enterprises This is the eBook version of the print title. The eBook edition contains the same content as the print edition. You will find instructions in the last few pages of your eBook that directs you to the media files. Business Intelligence Roadmap is a visual guide to developing an effective business intelligence (BI) decision-support application. This book outlines a methodology that takes into account the complexity of developing applications in an integrated BI environment. The authors walk readers through every step of the process--from strategic planning to the selection of new technologies and the evaluation of application releases. The book also serves as a single-source guide to the best practices of BI projects. Part I steers readers through the six stages of a BI project: justification, planning, business analysis, design, construction, and deployment. Each chapter describes one of sixteen development steps and the major activities, deliverables, roles, and responsibilities. All technical material is clearly expressed in tables, graphs, and diagrams. Part II provides five matrices that serve as references for the development process charted in Part I. Management tools, such as graphs illustrating the timing and coordination of activities, are included throughout the book. The authors conclude by crystallizing their many years of experience in a list of dos, don'ts, tips, and rules of thumb. Both the book and the methodology it describes are designed to adapt to the specific needs of individual stakeholders and organizations. The book directs business representatives, business sponsors, project managers, and technicians to the chapters that address their distinct responsibilities. The framework of the book allows organizations to begin at any step and enables projects to be scheduled and managed in a variety of ways. Business Intelligence Roadmap is a clear and comprehensive guide to negotiating the complexities inherent in the development of valuable business intelligence decision-support applications.

Data Lake for Enterprises - Tomcy John 2017-05-31

A practical guide to implementing your enterprise data lake using Lambda Architecture as the base About This Book Build a full-fledged data lake for your organization with popular big data technologies using the Lambda architecture as the base Delve into the big data technologies required to meet modern day business strategies A highly practical guide to implementing enterprise data lakes with lots of examples and real-world use-cases Who This Book Is For Java developers and architects who would like to implement a data lake for their enterprise will find this book useful. If you want to get hands-on experience with the Lambda Architecture and big data technologies by implementing a practical solution using these technologies, this book will also help you. What You Will Learn Build an enterprise-level data lake using the relevant big data technologies Understand the core of the Lambda architecture and how to apply it in an enterprise Learn the technical details around Sqoop and its functionalities Integrate Kafka with Hadoop components to acquire enterprise data Use flume with streaming technologies for stream-based processing Understand stream- based processing with reference to Apache Spark Streaming Incorporate Hadoop components and know the advantages they provide for enterprise data lakes Build fast, streaming, and high-performance applications using ElasticSearch Make your data ingestion process consistent across various data formats with configurability Process your data to derive intelligence using machine learning algorithms In Detail The term "Data Lake" has recently emerged as a prominent term in the big data industry. Data scientists can make use of it in deriving meaningful insights that can be used by businesses to redefine or transform the way they operate. Lambda architecture is also emerging as one of the very eminent patterns in the big data landscape, as it not only helps to derive useful information from historical

data but also correlates real-time data to enable business to take critical decisions. This book tries to bring these two important aspects — data lake and lambda architecture—together. This book is divided into three main sections. The first introduces you to the concept of data lakes, the importance of data lakes in enterprises, and getting you up-to-speed with the Lambda architecture. The second section delves into the principal components of building a data lake using the Lambda architecture. It introduces you to popular big data technologies such as Apache Hadoop, Spark, Sqoop, Flume, and ElasticSearch. The third section is a highly practical demonstration of putting it all together, and shows you how an enterprise data lake can be implemented, along with several real-world use-cases. It also shows you how other peripheral components can be added to the lake to make it more efficient. By the end of this book, you will be able to choose the right big data technologies using the lambda architectural patterns to build your enterprise data lake. Style and approach The book takes a pragmatic approach, showing ways to leverage big data technologies and lambda architecture to build an enterprise-level data lake.

Cloud Analytics with Microsoft Azure - Has Altaiar 2019-11-26

Leverage the power of Azure to get efficient data insights from your big data in real time Key Features Explore the basics of cloud analytics using Azure Discover different ways to process and visualize your data easily Learn to use Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse) to derive real-time customer insights Book Description With data being generated at an exponential speed, organizations all over the world are migrating their infrastructure to the cloud. Application management becomes much easier when you use a cloud platform to build, manage, and deploy your services and applications. Cloud Analytics with Microsoft Azure covers all that you need to extract useful insights from your data. You'll explore the power of data with big data analytics, the Internet of Things (IoT), machine learning, artificial intelligence, and DataOps. You'll also delve into data analytics by studying use cases that focus on creating actionable insights from near-real-time data. As you advance, you'll learn to build an end-to-end analytics pipeline on the cloud with machine learning and deep learning concepts. By the end of this book, you'll have developed a solid understanding of data analytics with Azure and its practical implementation. What you will learn Explore the concepts of modern data warehouses and data pipelines Discover different design considerations while applying a cloud analytics solution Design an end-to-end analytics pipeline on the cloud Differentiate between structured, semi-structured, and unstructured data Choose a cloud-based service for your data analytics solutions Use Azure services to ingest, store and analyze data of any scale Who this book is for If you're planning to adopt the cloud analytics model for your business, this book will help you understand the design and business considerations that you must keep in mind. Though not necessary, a basic understanding of data analytics concepts such as data streaming, data types, the machine learning life cycle, and Docker containers will help you get the most out of the book.

The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights -

Robert Laberge 2011-05-12

Develop a custom, agile data warehousing and business intelligence architecture Empower your users and drive better decision making across your enterprise with detailed instructions and best practices from an expert developer and trainer. The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights shows how to plan, design, construct, and administer an integrated end-to-end DW/BI solution. Learn how to choose appropriate components, build an enterprise data model, configure data marts and data warehouses, establish data flow, and mitigate risk. Change management, data governance, and security are also covered in this comprehensive guide. Understand the components of BI and data warehouse systems Establish project goals and implement an effective deployment plan Build accurate logical and physical enterprise data models Gain insight into your company's transactions with data mining Input, cleanse, and normalize data using ETL (Extract, Transform, and Load) techniques Use structured input files to define data requirements Employ top-down, bottom-up, and hybrid design methodologies Handle security and optimize performance using data governance tools Robert Laberge is the founder of several Internet ventures and a principle consultant for the IBM Industry Models and Assets Lab, which has a focus on data warehousing and business intelligence solutions.

Data Governance - John Ladley 2019-11-08

Managing data continues to grow as a necessity for modern organizations. There are seemingly infinite

opportunities for organic growth, reduction of costs, and creation of new products and services. It has become apparent that none of these opportunities can happen smoothly without data governance. The cost of exponential data growth and privacy / security concerns are becoming burdensome. Organizations will encounter unexpected consequences in new sources of risk. The solution to these challenges is also data governance; ensuring balance between risk and opportunity. Data Governance, Second Edition, is for any executive, manager or data professional who needs to understand or implement a data governance program. It is required to ensure consistent, accurate and reliable data across their organization. This book offers an overview of why data governance is needed, how to design, initiate, and execute a program and how to keep the program sustainable. This valuable resource provides comprehensive guidance to beginning professionals, managers or analysts looking to improve their processes, and advanced students in Data Management and related courses. With the provided framework and case studies all professionals in the data governance field will gain key insights into launching successful and money-saving data governance program. Incorporates industry changes, lessons learned and new approaches Explores various ways in which data analysts and managers can ensure consistent, accurate and reliable data across their organizations Includes new case studies which detail real-world situations Explores all of the capabilities an organization must adopt to become data driven Provides guidance on various approaches to data governance, to determine whether an organization should be low profile, central controlled, agile, or traditional Provides guidance on using technology and separating vendor hype from sincere delivery of necessary capabilities Offers readers insights into how their organizations can improve the value of their data, through data quality, data strategy and data literacy Provides up to 75% brand-new content compared to the first edition

A Practical Guide to Managing Reference Data with IBM InfoSphere Master Data Management

Reference Data Management Hub - Whei-Jen Chen 2013-05-06

IBM® InfoSphere® Master Data Management Reference Data Management Hub (InfoSphere MDM Ref DM Hub) is designed as a ready-to-run application that provides the governance, process, security, and audit control for managing reference data as an enterprise standard, resulting in fewer errors, reduced business risk and cost savings. This IBM Redbooks® publication describes where InfoSphere MDM Ref DM Hub fits into information management reference architecture. It explains the end-to-end process of an InfoSphere MDM Ref DM Hub implementation including the considerations of planning a reference data management project, requirements gathering and analysis, model design in detail, and integration considerations and scenarios. It then shows implementation examples and the ongoing administration tasks. This publication can help IT professionals who are interested or have a need to manage reference data efficiently and implement an InfoSphere MDM Ref DM Hub solution with ease.

Population Health Analytics - Martha L. Sylvia 2021-03

"Binding: PB"--

Data Quality - Carlo Batini 2006-09-27

Poor data quality can seriously hinder or damage the efficiency and effectiveness of organizations and businesses. The growing awareness of such repercussions has led to major public initiatives like the "Data Quality Act" in the USA and the "European 2003/98" directive of the European Parliament. Batini and Scannapieco present a comprehensive and systematic introduction to the wide set of issues related to data quality. They start with a detailed description of different data quality dimensions, like accuracy, completeness, and consistency, and their importance in different types of data, like federated data, web data, or time-dependent data, and in different data categories classified according to frequency of change, like stable, long-term, and frequently changing data. The book's extensive description of techniques and methodologies from core data quality research as well as from related fields like data mining, probability theory, statistical data analysis, and machine learning gives an excellent overview of the current state of the art. The presentation is completed by a short description and critical comparison of tools and practical methodologies, which will help readers to resolve their own quality problems. This book is an ideal combination of the soundness of theoretical foundations and the applicability of practical approaches. It is ideally suited for everyone - researchers, students, or professionals - interested in a comprehensive overview of data quality issues. In addition, it will serve as the basis for an introductory course or for self-

study on this topic.

Data Modeling Essentials - Graeme Simsion 2004-12-03

Data Modeling Essentials, Third Edition, covers the basics of data modeling while focusing on developing a facility in techniques, rather than a simple familiarization with "the rules". In order to enable students to apply the basics of data modeling to real models, the book addresses the realities of developing systems in real-world situations by assessing the merits of a variety of possible solutions as well as using language and diagramming methods that represent industry practice. This revised edition has been given significantly expanded coverage and reorganized for greater reader comprehension even as it retains its distinctive hallmarks of readability and usefulness. Beginning with the basics, the book provides a thorough grounding in theory before guiding the reader through the various stages of applied data modeling and database design. Later chapters address advanced subjects, including business rules, data warehousing, enterprise-wide modeling and data management. It includes an entirely new section discussing the development of logical and physical modeling, along with new material describing a powerful technique for model verification. It also provides an excellent resource for additional lectures and exercises. This text is the ideal reference for data modelers, data architects, database designers, DBAs, and systems analysts, as well as undergraduate and graduate-level students looking for a real-world perspective. Thorough coverage of the fundamentals and relevant theory. Recognition and support for the creative side of the process. Expanded coverage of applied data modeling includes new chapters on logical and physical database design. New material describing a powerful technique for model verification. Unique coverage of the practical and human aspects of modeling, such as working with business specialists, managing change, and resolving conflict.

Learning SQL - Alan Beaulieu 2009-04-11

Updated for the latest database management systems -- including MySQL 6.0, Oracle 11g, and Microsoft's SQL Server 2008 -- this introductory guide will get you up and running with SQL quickly. Whether you need to write database applications, perform administrative tasks, or generate reports, Learning SQL, Second Edition, will help you easily master all the SQL fundamentals. Each chapter presents a self-contained lesson on a key SQL concept or technique, with numerous illustrations and annotated examples. Exercises at the end of each chapter let you practice the skills you learn. With this book, you will: Move quickly through SQL basics and learn several advanced features Use SQL data statements to generate, manipulate, and retrieve data Create database objects, such as tables, indexes, and constraints, using SQL schema statements Learn how data sets interact with queries, and understand the importance of subqueries Convert and manipulate data with SQL's built-in functions, and use conditional logic in data statements Knowledge of SQL is a must for interacting with data. With Learning SQL, you'll quickly learn how to put the power and flexibility of this language to work.

Data Mining and Data Warehousing - Parteek Bhatia 2019-04-30

Written in lucid language, this valuable textbook brings together fundamental concepts of data mining and data warehousing in a single volume. Important topics including information theory, decision tree, Naïve Bayes classifier, distance metrics, partitioning clustering, associate mining, data marts and operational data store are discussed comprehensively. The textbook is written to cater to the needs of undergraduate students of computer science, engineering and information technology for a course on data mining and data warehousing. The text simplifies the understanding of the concepts through exercises and practical examples. Chapters such as classification, associate mining and cluster analysis are discussed in detail with their practical implementation using Weka and R language data mining tools. Advanced topics including big data analytics, relational data models and NoSQL are discussed in detail. Pedagogical features including unsolved problems and multiple-choice questions are interspersed throughout the book for better understanding.

Agile Data Warehouse Design - Lawrence Corr 2011-04-16

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

Lawrence Corr 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.

The Biml Book - Andy Leonard 2017-10-30

Learn Business Intelligence Markup Language (Biml) for automating much of the repetitive, manual labor involved in data integration. We teach you how to build frameworks and use advanced Biml features to get more out of SQL Server Integration Services (SSIS), Transact-SQL (T-SQL), and SQL Server Analysis Services (SSAS) than you ever thought possible. The first part of the book starts with the basics—getting your development environment configured, Biml syntax, and scripting essentials. Whether a beginner or a seasoned Biml expert, the next part of the book guides you through the process of using Biml to build a framework that captures both your design patterns and execution management. Design patterns are reusable code blocks that standardize the approach you use to perform certain types of data integration, logging, and other key data functions. Design patterns solve common problems encountered when developing data integration solutions. Because you do not have to build the code from scratch each time, design patterns improve your efficiency as a Biml developer. In addition to leveraging design patterns in your framework, you will learn how to build a robust metadata store and how to package your framework into Biml bundles for deployment within your enterprise. In the last part of the book, we teach you more advanced Biml features and capabilities, such as SSAS development, T-SQL recipes, documentation autogeneration, and Biml troubleshooting. The Biml Book: Provides practical and applicable examples Teaches you how to use Biml to reduce development time while improving quality Takes you through solutions to common data integration and BI challenges What You'll Learn Master the basics of Business Intelligence Markup Language (Biml) Study patterns for automating SSIS package generation Build a Biml Framework Import and transform database schemas Automate generation of scripts and projects Who This Book Is For BI developers wishing to quickly locate previously tested solutions, Microsoft BI specialists, those seeking more information about solution automation and code generation, and practitioners of Data Integration Lifecycle Management (DILM) in the DevOps enterprise

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

The Analytical Puzzle - David Haertzen 2012-07-01

Do you enjoy completing puzzles? Perhaps one of the most challenging (yet rewarding) puzzles is delivering a successful data warehouse suitable for data mining and analytics. The Analytical Puzzle describes an unbiased, practical, and comprehensive approach to building a data warehouse which will lead to an increased level of business intelligence within your organization. New technologies continuously impact this approach and therefore this book explains how to leverage big data, cloud computing, data warehouse appliances, data mining, predictive analytics, data visualization and mobile devices. Here are the main objectives for each of the book's 19 chapters:

- Chapter 1: Develop a foundational knowledge of data warehousing, business intelligence and analytics
- Chapter 2: Build the business case needed to sell your data warehousing project, and then produce a project plan that avoids common pitfalls
- Chapter 3: Elicit and organize business intelligence and data warehousing business requirements
- Chapter 4: Specify the technical architecture of the data warehousing system, including software and infrastructure components, technology stack, and non-functional requirements. Gain an understanding of cloud based data warehousing and data warehouse appliances
- Chapter 5: Learn about data attributes including metrics and key performance indicators (KPIs), the raw material of data warehousing and business intelligence
- Chapter 6: Learn about data modeling and how to apply design patterns for each part of the data warehouse
- Chapter 7: Speak the dimensional modeling language of measures, dimensions, facts, cubes, stars, and snowflakes
- Chapter 8: Organize a successful data governance program. Learn how to manage metadata for your data warehousing and business intelligence project
- Chapter 9: Identify useful data sources and implement a data quality program
- Chapter 10: Use database technology for your data warehousing project, and understand the impact of data warehouse appliances, big data, in memory databases, columnar databases and OnLine Analytical Processing (OLAP)
- Chapter 11: Apply data integration and understand the role data mapping, data cleansing, data transformation, and loading data play in a successful data warehouse
- Chapter 12: Use the business intelligence (BI) operations of slice, dice, drill down, roll up, and pivot to analyze and present data
- Chapter 13: Learn about descriptive and predictive statistics, and calculate mean, median, mode, variance and standard deviation
- Chapter 14: Harness analytical methods such as regression analysis, data mining, and statistics to make profitable decisions and anticipate the future
- Chapter 15: Appreciate the components and design patterns that compose a successful analytic application
- Chapter 16: Gain an understanding of the uses and benefits of scorecards and dashboards including support of mobile device users
- Chapter 17: Gain insight into applications of business intelligence that could profit your organization, including risk management, finance, marketing, government, healthcare, science and sports
- Chapter 18: Perform customer analytics to better understand and segment your customers
- Chapter 19: Test, roll out, and sustain the data warehouse

Microsoft SQL Server 2014 Business Intelligence Development Beginner's Guide - Reza Rad 2014-05-18

Written in an easy-to-follow, example-driven format, there are plenty of step-by-step instructions to help get you started! The book has a friendly approach, with the opportunity to learn by experimenting. If you are a BI and Data Warehouse developer new to Microsoft Business Intelligence, and looking to get a good understanding of the different components of Microsoft SQL Server for Business Intelligence, this book is for you. It's assumed that you will have some experience in databases systems and T-SQL. This book is will

give you a good upshot view of each component and scenarios featuring th.

Data Warehousing and Analytics - David Taniar 2022-02-04

This textbook covers all central activities of data warehousing and analytics, including transformation, preparation, aggregation, integration, and analysis. It discusses the full spectrum of the journey of data from operational/transactional databases, to data warehouses and data analytics; as well as the role that data warehousing plays in the data processing lifecycle. It also explains in detail how data warehouses may be used by data engines, such as BI tools and analytics algorithms to produce reports, dashboards, patterns, and other useful information and knowledge. The book is divided into six parts, ranging from the basics of data warehouse design (Part I - Star Schema, Part II - Snowflake and Bridge Tables, Part III - Advanced Dimensions, and Part IV - Multi-Fact and Multi-Input), to more advanced data warehousing concepts (Part V - Data Warehousing and Evolution) and data analytics (Part VI - OLAP, BI, and Analytics). This textbook approaches data warehousing from the case study angle. Each chapter presents one or more case studies to thoroughly explain the concepts and has different levels of difficulty, hence learning is incremental. In addition, every chapter has also a section on further readings which give pointers and references to research papers related to the chapter. All these features make the book ideally suited for either introductory courses on data warehousing and data analytics, or even for self-studies by professionals. The book is accompanied by a web page that includes all the used datasets and codes as well as slides and solutions to exercises.

Expert Cube Development with SSAS Multidimensional Models - Chris Webb 2014-02-24

An easy-to-follow guide full of hands on examples of real-world Analysis Services cube development tasks. Each topic is explained and placed in context, and for the more inquisitive reader, there also more in-depth details of the concepts used. If you are an Analysis Services cube designer wishing to learn more advanced topic and best practices for cube design, this book is for you. You are expected to have some prior experience with Analysis Services cube development.

Database Systems - Nenad Jukic 2013-01-03

An introductory, yet comprehensive, database textbook intended for use in undergraduate and graduate information systems database courses. This text also provides practical content to current and aspiring information systems, business data analysis, and decision support industry professionals. Database Systems: Introduction to Databases and Data Warehouses covers both analytical and operations database as knowledge of both is integral to being successful in today's business environment. It also provides a solid theoretical foundation and hands-on practice using an integrated web-based data-modeling suite.

Professional Microsoft SQL Server 2008 Integration Services - Brian Knight 2011-01-31

The new edition of the successful previous version is 25 percent revised and packed with more than 200 pages of new material on the 2008 release of SQL Server Integration Services (SSIS) Renowned author Brian Knight and his expert coauthors show developers how to master the 2008 release of SSIS, which is both more powerful and more complex than ever Case studies and tutorial examples acquired over the three years since the previous edition will contribute to helping illustrate advanced concepts and techniques New chapters include coverage of data warehousing using SSIS, new methods for managing the SSIS platform, and improved techniques for ETL operations

The Data Catalog: Sherlock Holmes Data Sleuthing for Analytics - Bonnie K. O'Neil 2020-03-16

Apply this definitive guide to data catalogs and select the feature set needed to empower your data citizens in their quest for faster time to insight. The data catalog may be the most important breakthrough in data management in the last decade, ranking alongside the advent of the data warehouse. The latter enabled business consumers to conduct their own analyses to obtain insights themselves. The data catalog is the next wave of this, empowering business users even further to drastically reduce time to insight, despite the rising tide of data flooding the enterprise. Use this book as a guide to provide a broad overview of the most popular Machine Learning (ML) data catalog products, and perform due diligence using the extensive features list. Consider graphical user interface (GUI) design issues such as layout and navigation, as well as scalability in terms of how the catalog will handle your current and anticipated data and metadata needs. O'Neil & Fryman...present a typology which ranges from products that focus on data lineage, curation and search, data governance, data preparation, and of course, the core capability of finding and understanding

the data. The authors emphasize that machine learning is being adopted in many of these products, enabling a more elegant data democratization solution in the face of the burgeoning mountain of data that is engulfing organizations. Derek Strauss, Chairman/CEO, Gavroshe, and Former CDO, TD Ameritrade This book is organized into three sections: · Chapters 1 and 2 reveal the rationale for a data catalog and share how data scientists, data administrators, and curators fare with and without a data catalog. · Chapters 3-10 present the many different types of data catalogs. · Chapters 11 and 12 provide an extensive features' list, current trends, and visions for the future.

Super Charge Your Data Warehouse - Dan Linstedt 2011-11-01

Do You Know If Your Data Warehouse Flexible, Scalable, Secure and Will It Stand The Test Of Time And Avoid Being Part Of The Dreaded "Life Cycle"? The Data Vault took the Data Warehouse world by storm when it was released in 2001. Some of the world's largest and most complex data warehouse situations understood the value it gave especially with the capabilities of unlimited scaling, flexibility and security. Here is what industry leaders say about the Data Vault "The Data Vault is the optimal choice for modeling the EDW in the DW 2.0 framework" - Bill Inmon, The Father of Data Warehousing "The Data Vault is foundationally strong and an exceptionally scalable architecture" - Stephen Brobst, CTO, Teradata "The Data Vault should be considered as a potential standard for RDBMS-based analytic data management by organizations looking to achieve a high degree of flexibility, performance and openness" - Doug Laney, Deloitte Analytics Institute "I applaud Dan's contribution to the body of Business Intelligence and Data Warehousing knowledge and recommend this book be read by both data professionals and end users" - Howard Dresner, From the Foreword - Speaker, Author, Leading Research Analyst and Advisor You have in your hands the work, experience and testing of 2 decades of building data warehouses. The Data Vault model and methodology has proven itself in hundreds (perhaps thousands) of solutions in Insurance, Crime-Fighting, Defense, Retail, Finance, Banking, Power, Energy, Education, High-Tech and many more. Learn the techniques and implement them and learn how to build your Data Warehouse faster than you have ever done before while designing it to grow and scale no matter what you throw at it. Ready to "Super Charge Your Data Warehouse"?

Agile Analytics - Ken Collier 2012

Using Agile methods, you can bring far greater innovation, value, and quality to any data warehousing (DW), business intelligence (BI), or analytics project. However, conventional Agile methods must be carefully adapted to address the unique characteristics of DW/BI projects. In Agile Analytics, Agile pioneer Ken Collier shows how to do just that. Collier introduces platform-agnostic Agile solutions for integrating infrastructures consisting of diverse operational, legacy, and specialty systems that mix commercial and custom code. Using working examples, he shows how to manage analytics development teams with widely diverse skill sets and how to support enormous and fast-growing data volumes. Collier's techniques offer optimal value whether your projects involve "back-end" data management, "front-end" business analysis, or both. Part I focuses on Agile project management techniques and delivery team coordination, introducing core practices that shape the way your Agile DW/BI project community can collaborate toward success Part II presents technical methods for enabling continuous delivery of business value at production-quality levels, including evolving superior designs; test-driven DW development; version control; and project automation Collier brings together proven solutions you can apply right now--whether you're an IT decision-maker, data warehouse professional, database administrator, business intelligence specialist, or database developer. With his help, you can mitigate project risk, improve business alignment, achieve better results--and have fun along the way.

Indian National Bibliography - B. S. Kesavan 2012-05

The Book of R - Tilman M. Davies 2016-07-16

The Book of R is a comprehensive, beginner-friendly guide to R, the world's most popular programming language for statistical analysis. Even if you have no programming experience and little more than a grounding in the basics of mathematics, you'll find everything you need to begin using R effectively for statistical analysis. You'll start with the basics, like how to handle data and write simple programs, before moving on to more advanced topics, like producing statistical summaries of your data and performing

statistical tests and modeling. You'll even learn how to create impressive data visualizations with R's basic graphics tools and contributed packages, like ggplot2 and ggvis, as well as interactive 3D visualizations using the rgl package. Dozens of hands-on exercises (with downloadable solutions) take you from theory to practice, as you learn: -The fundamentals of programming in R, including how to write data frames, create functions, and use variables, statements, and loops -Statistical concepts like exploratory data analysis, probabilities, hypothesis tests, and regression modeling, and how to execute them in R -How to access R's thousands of functions, libraries, and data sets -How to draw valid and useful conclusions from your data -How to create publication-quality graphics of your results Combining detailed explanations with real-world examples and exercises, this book will provide you with a solid understanding of both statistics and the depth of R's functionality. Make The Book of R your doorway into the growing world of data analysis.

Data Warehousing - 2001-04-27

Rapid access to information is a prime requirement in any organization that wants to have a competitive edge in today's fast changing markets. How to retrieve information? How to capture data? How to format it? The answer lies in Data Warehousing. This HOTT Guide will give you access to all the essential information about the newest data storehouse: through articles by expert trendwachers on strategic considerations, how-to reports defining the various ways to extract the data needed for critical business decisions, technical papers clarifying technologies and tools, business cases and key concepts that will provide the reader with a comprehensive overview of a business solution that is already indispensable.

Patterns of Information Management - Mandy Chessell 2013

In the era of "Big Data," effective information management has become crucial to the success of virtually all organizations. Unfortunately, few IT practitioners know today's best practices for successfully managing enterprise information resources. Patterns for Information Management offers the solution: a multi-disciplinary patterns-based approach that reflects where information comes from, how it is distributed, protected, governed, monitored -- and, ultimately, utilized.

SQL Server Integration Services Design Patterns - Tim Mitchell 2014-12-24

SQL Server Integration Services Design Patterns is newly-revised for SQL Server 2014, and is a book of recipes for SQL Server Integration Services (SSIS). Design patterns in the book help to solve common problems encountered when developing data integration solutions. The patterns and solution examples in the book increase your efficiency as an SSIS developer, because you do not have to design and code from scratch with each new problem you face. The book's team of expert authors take you through numerous design patterns that you'll soon be using every day, providing the thought process and technical details needed to support their solutions. SQL Server Integration Services Design Patterns goes beyond the surface of the immediate problems to be solved, delving into why particular problems should be solved in

certain ways. You'll learn more about SSIS as a result, and you'll learn by practical example. Where appropriate, the book provides examples of alternative patterns and discusses when and where they should be used. Highlights of the book include sections on ETL Instrumentation, SSIS Frameworks, Business Intelligence Markup Language, and Dependency Services. Takes you through solutions to common data integration challenges Provides examples involving Business Intelligence Markup Language Teaches SSIS using practical examples

Practical Data Analysis Using Jupyter Notebook - Marc Wintjen 2020-06-19

Understand data analysis concepts to make accurate decisions based on data using Python programming and Jupyter Notebook Key FeaturesFind out how to use Python code to extract insights from data using real-world examplesWork with structured data and free text sources to answer questions and add value using dataPerform data analysis from scratch with the help of clear explanations for cleaning, transforming, and visualizing dataBook Description Data literacy is the ability to read, analyze, work with, and argue using data. Data analysis is the process of cleaning and modeling your data to discover useful information. This book combines these two concepts by sharing proven techniques and hands-on examples so that you can learn how to communicate effectively using data. After introducing you to the basics of data analysis using Jupyter Notebook and Python, the book will take you through the fundamentals of data. Packed with practical examples, this guide will teach you how to clean, wrangle, analyze, and visualize data to gain useful insights, and you'll discover how to answer questions using data with easy-to-follow steps. Later chapters teach you about storytelling with data using charts, such as histograms and scatter plots. As you advance, you'll understand how to work with unstructured data using natural language processing (NLP) techniques to perform sentiment analysis. All the knowledge you gain will help you discover key patterns and trends in data using real-world examples. In addition to this, you will learn how to handle data of varying complexity to perform efficient data analysis using modern Python libraries. By the end of this book, you'll have gained the practical skills you need to analyze data with confidence. What you will learnUnderstand the importance of data literacy and how to communicate effectively using dataFind out how to use Python packages such as NumPy, pandas, Matplotlib, and the Natural Language Toolkit (NLTK) for data analysisWrangle data and create DataFrames using pandasProduce charts and data visualizations using time-series datasetsDiscover relationships and how to join data together using SQLUse NLP techniques to work with unstructured data to create sentiment analysis modelsDiscover patterns in real-world datasets that provide accurate insightsWho this book is for This book is for aspiring data analysts and data scientists looking for hands-on tutorials and real-world examples to understand data analysis concepts using SQL, Python, and Jupyter Notebook. Anyone looking to evolve their skills to become data-driven personally and professionally will also find this book useful. No prior knowledge of data analysis or programming is required to get started with this book.