Apache Spark In 24 Hours Sams Teach Yourself Sams Teach Yourself In 24 Hours

This is likewise one of the factors by obtaining the soft documents of this **Apache Spark In 24 Hours Sams Teach Yourself Sams Teach Yourself In 24 Hours** by online. You might not require more times to spend to go to the books establishment as skillfully as search for them. In some cases, you likewise accomplish not discover the statement Apache Spark In 24 Hours Sams Teach Yourself Sams Teach Yourself In 24 Hours that you are looking for. It will completely squander the time.

However below, subsequently you visit this web page, it will be consequently agreed simple to get as with ease as download guide Apache Spark In 24 Hours Sams Teach Yourself Sams Teach Yourself In 24 Hours

It will not acknowledge many grow old as we explain before. You can get it even though perform something else at house and even in your workplace. so easy! So, are you question? Just exercise just what we give below as well as review **Apache Spark In 24 Hours Sams Teach Yourself Sams Teach Yourself In 24 Hours** what you behind to read!

Data Analytics with Hadoop - Benjamin Bengfort 2016-06 Ready to use statistical and machine-learning techniques across large data sets? This practical guide shows you why the Hadoop ecosystem is perfect for the job. Instead of

deployment, operations, or software development usually associated with distributed computing, you'll focus on particular analyses you can build, the data warehousing techniques that Hadoop provides, and higher order data workflows this framework can produce. Data scientists and analysts will learn how to perform a wide range of techniques, from writing MapReduce and Spark applications with Python to using advanced modeling and data management with Spark MLlib, Hive, and HBase. You'll also learn about the analytical processes and data systems available to build and empower data products that can handle—and actually require—huge amounts of data. Understand core concepts behind Hadoop and cluster computing Use design patterns and parallel analytical algorithms to create distributed data analysis jobs Learn about data management, mining, and warehousing in a distributed context using Apache Hive and HBase Use Sqoop and Apache Flume to ingest data from relational databases Program complex Hadoop and Spark applications with Apache Pig and Spark DataFrames Perform machine learning techniques such as classification, clustering, and collaborative filtering with Spark's MLlib

Learning Spark - Jules S. Damji 2020-07-16
Data is bigger, arrives faster, and comes in a variety of formats—and it all needs to be processed at scale for analytics or machine learning. But how can you process such varied workloads efficiently? Enter Apache Spark. Updated to include Spark 3.0, this second edition shows data engineers and data scientists why structure and unification in Spark matters. Specifically, this book explains how to perform simple and complex data analytics and employ machine learning algorithms. Through step-by-step walk-throughs, code snippets, and notebooks, you'll be able to: Learn Python, SQL, Scala, or Java high-level Structured APIs

Understand Spark operations and SQL Engine Inspect, tune, and debug Spark operations with Spark configurations and Spark UI Connect to data sources: JSON, Parquet, CSV, Avro, ORC, Hive, S3, or Kafka Perform analytics on batch and streaming data using Structured Streaming Build reliable data pipelines with open source Delta Lake and Spark Develop machine learning pipelines with MLlib and productionize models using MLflow

Hadoop For Dummies - Dirk deRoos 2014-04-14
Let Hadoop For Dummies help harness the power of your data and rein in the information overload Big data has become big business, and companies and organizations of all sizes are struggling to find ways to retrieve valuable information from their massive data sets with becoming overwhelmed. Enter Hadoop and this easy-to-understand For Dummies guide. Hadoop For Dummies helps readers understand the value of big data, make a business case for using Hadoop, navigate the Hadoop ecosystem, and build and manage Hadoop applications and clusters. Explains the origins of Hadoop, its economic benefits, and its functionality and practical applications Helps you find your way around the Hadoop ecosystem, program MapReduce, utilize design patterns, and get your Hadoop cluster up and running quickly and easily Details how to use Hadoop applications for data mining, web analytics and personalization, large-scale text processing, data science, and problem-solving Shows you how to improve the value of your Hadoop cluster, maximize your investment in Hadoop, and avoid common pitfalls when building your Hadoop cluster From programmers challenged with building and maintaining affordable, scaleable data systems to administrators who must deal with huge volumes of information effectively and efficiently, this how-to has something to help you with Hadoop.
Advanced Analytics with Spark

- Sandy Ryza 2015-04-02 In this practical book, four Cloudera data scientists present a set of self-contained patterns for performing large-scale data analysis with Spark. The authors bring Spark, statistical methods, and real-world data sets together to teach you how to approach analytics problems by example. You'll start with an introduction to Spark and its ecosystem, and then dive into patterns that apply common techniques—classification, collaborative filtering, and anomaly detection among others—to fields such as genomics, security, and finance. If you have an entry-level understanding of machine learning and statistics, and you program in Java, Python, or Scala, you'll find these patterns useful for working on your own data applications. Patterns include: Recommending music and the Audioscrobbler data set Predicting forest cover with decision trees Anomaly detection in network traffic with K-means clustering Understanding Wikipedia with Latent Semantic Analysis Analyzing co-occurrence networks with GraphX Geospatial and temporal data analysis on the New York City Taxi Trips data Estimating financial risk through Monte Carlo simulation Analyzing genomics data and the BDG project Analyzing neuroimaging data with PySpark and Thunder

Hadoop: The Definitive Guide - Tom White 2012-05-10 Ready to unlock the power of your data? With this comprehensive guide, you'll learn how to build and maintain reliable, scalable, distributed systems with Apache Hadoop. This book is ideal for programmers looking to analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters. You'll find illuminating case studies that demonstrate how Hadoop is used to solve specific problems. This third edition covers recent changes to Hadoop, including material on the new MapReduce API, as well as MapReduce 2 and its more

flexible execution model (YARN). Store large datasets with the Hadoop Distributed File System (HDFS) Run distributed computations with MapReduce Use Hadoop's data and I/O building blocks for compression, data integrity, serialization (including Avro), and persistence Discover common pitfalls and advanced features for writing real-world MapReduce programs Design, build, and administer a dedicated Hadoop cluster—or run Hadoop in the cloud Load data from relational databases into HDFS, using Sqoop Perform large-scale data processing with the Pig query language Analyze datasets with Hive, Hadoop's data warehousing system Take advantage of HBase for structured and semi-structured data, and ZooKeeper for building distributed systems Programming Hive - Edward Capriolo 2012-09-26 Describes the features and functions of Apache Hive, the data infrastructure for Hadoop. **Hadoop in 24 Hours, Sams Teach Yourself** - Jeffrey Aven

2017-01-06
Apache Hadoop is the technology at the heart of the Big Data revolution, and Hadoop skills are in enormous demand. Now, in just 24 lessons of one hour or less, you can learn all the skills and techniques you'll need to deploy each key component of a Hadoop platform in your local environment or in the cloud, building a fully functional Hadoop cluster and using it with real programs and datasets. Each short, easy lesson builds on all that's come before, helping you master all of Hadoop's essentials, and extend it to meet your unique challenges. Apache Hadoop in 24 Hours, Sams Teach Yourself covers all this, and much more: Understanding Hadoop and the Hadoop Distributed File System (HDFS) Importing data into Hadoop, and process it there Mastering basic MapReduce Java programming, and using advanced MapReduce API concepts Making the most of Apache Pig and Apache Hive Implementing and administering YARN

Taking advantage of the full Hadoop ecosystem Managing Hadoop clusters with Apache Ambari Working with the Hadoop User Environment (HUE) Scaling, securing, and troubleshooting Hadoop environments Integrating Hadoop into the enterprise Deploying Hadoop in the cloud Getting started with Apache Spark Step-by-step instructions walk you through common questions, issues, and tasks; Q-and-As, Quizzes, and Exercises build and test your knowledge; "Did You Know?" tips offer insider advice and shortcuts; and "Watch Out!" alerts help you avoid pitfalls. By the time you're finished, you'll be comfortable using Apache Hadoop to solve a wide spectrum of Big Data problems.
Database Reliability Engineering - Laine Campbell 2017-10-26
The infrastructure-as-code revolution in IT is also affecting database administration. With this practical book, developers, system administrators, and junior to mid-level DBAs will learn how the modern practice of site reliability engineering applies to the craft of database architecture and operations. Authors Laine Campbell and Charity Majors provide a framework for professionals looking to join the ranks of today's database reliability engineers (DBRE). You'll begin by exploring core operational concepts that DBREs need to master. Then you'll examine a wide range of database persistence options, including how to implement key technologies to provide resilient, scalable, and performant data storage and retrieval. With a firm foundation in database reliability engineering, you'll be ready to dive into the architecture and operations of any modern database. This book covers: Service-level requirements and risk management Building and evolving an architecture for operational visibility Infrastructure engineering and infrastructure management How to facilitate the release management process Data storage, indexing, and

replication Identifying datastore characteristics and best use cases Datastore architectural components and data-driven architectures *Efficient R Programming* - Colin Gillespie 2016-12-08 There are many excellent R resources for visualization, data science, and package development. Hundreds of scattered vignettes, web pages, and forums explain how to use R in particular domains. But little has been written on how to simply make R work effectively—until now. This hands-on book teaches novices and experienced R users how to write efficient R code. Drawing on years of experience teaching R courses, authors Colin Gillespie and Robin Lovelace provide practical advice on a range of topics—from optimizing the set-up of RStudio to leveraging C++—that make this book a useful addition to any R user's bookshelf. Academics, business users, and programmers from a wide range of backgrounds stand to benefit from the guidance in Efficient R Programming. Get advice for setting up an R programming environment Explore general programming concepts and R coding techniques Understand the ingredients of an efficient R workflow Learn how to efficiently read and write data in R Dive into data carpentry—the vital skill for cleaning raw data Optimize your code with profiling, standard tricks, and other methods Determine your hardware capabilities for handling R computation Maximize the benefits of collaborative R programming Accelerate your transition from R hacker to R programmer **Spark in Action** - Jean-Georges Perrin 2020-05-12 Summary The Spark distributed data processing platform provides an easy-to-implement tool for ingesting, streaming, and processing data from any source. In Spark in Action, Second Edition, you'll learn to take advantage of Spark's core features and incredible processing speed, with applications including real-time computation, delayed

evaluation, and machine learning. Spark skills are a hot commodity in enterprises worldwide, and with Spark's powerful and flexible Java APIs, you can reap all the benefits without first learning Scala or Hadoop. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Analyzing enterprise data starts by reading, filtering, and merging files and streams from many sources. The Spark data processing engine handles this varied volume like a champ, delivering speeds 100 times faster than Hadoop systems. Thanks to SQL support, an intuitive interface, and a straightforward multilanguage API, you can use Spark without learning a complex new ecosystem. About the book Spark in Action, Second Edition, teaches you to create end-to-end analytics applications. In this entirely new book, you'll learn from interesting Java-based examples, including a complete data pipeline for processing NASA satellite data. And you'll discover Java, Python, and Scala code samples hosted on GitHub that you can explore and adapt, plus appendixes that give you a cheat sheet for installing tools and understanding Spark-specific terms. What's inside Writing Spark applications in Java Spark application architecture Ingestion through files, databases, streaming, and Elasticsearch Querying distributed datasets with Spark SQL About the reader This book does not assume previous experience with Spark, Scala, or Hadoop. About the author Jean-Georges Perrin is an experienced data and software architect. He is France's first IBM Champion and has been honored for 12 consecutive years. Table of Contents PART 1 - THE THEORY CRIPPLED BY AWESOME EXAMPLES 1 So, what is Spark, anyway? 2 Architecture and flow 3 The majestic role of the dataframe 4 Fundamentally lazy 5 Building a simple app for deployment 6 Deploying your simple app PART 2 -

INGESTION 7 Ingestion from files 8 Ingestion from databases 9 Advanced ingestion: finding data sources and building your own 10 Ingestion through structured streaming PART 3 - TRANSFORMING YOUR DATA 11 Working with SQL 12 Transforming your data 13 Transforming entire documents 14 Extending transformations with user-defined functions 15 Aggregating your data PART 4 - GOING FURTHER 16 Cache and checkpoint: Enhancing Spark's performances 17 Exporting data and building full data pipelines 18 Exploring deployment

Kafka: The Definitive Guide - Neha Narkhede 2017-08-31 Every enterprise application creates data, whether it's log messages, metrics, user activity, outgoing messages, or something else. And how to move all of this data becomes nearly as important as the data itself. If you're an application architect, developer, or production engineer new to Apache Kafka, this practical guide shows you how to use this open source streaming platform to handle real-time data feeds. Engineers from Confluent and LinkedIn who are responsible for developing Kafka explain how to deploy production Kafka clusters, write reliable event-driven microservices, and build scalable stream-processing applications with this platform. Through detailed examples, you'll learn Kafka's design principles, reliability guarantees, key APIs, and architecture details, including the replication protocol, the controller, and the storage layer. Understand publish-subscribe messaging and how it fits in the big data ecosystem. Explore Kafka producers and consumers for writing and reading messages Understand Kafka patterns and use-case requirements to ensure reliable data delivery Get best practices for building data pipelines and applications with Kafka Manage Kafka in production, and learn to perform monitoring, tuning, and maintenance tasks Learn the most critical metrics among

Kafka's operational measurements Explore how Kafka's stream delivery capabilities make it a perfect source for stream processing systems

Data Algorithms - Mahmoud Parsian 2015-07-13 If you are ready to dive into the MapReduce framework for processing large datasets, this practical book takes you step by step through the algorithms and tools you need to build distributed MapReduce applications with Apache Hadoop or Apache Spark. Each chapter provides a recipe for solving a massive computational problem, such as building a recommendation system. You'll learn how to implement the appropriate MapReduce solution with code that you can use in your projects. Dr. Mahmoud Parsian covers basic design patterns, optimization techniques, and data mining and machine learning solutions for problems in bioinformatics, genomics, statistics, and social network analysis. This book also includes an overview of MapReduce, Hadoop, and Spark. Topics include: Market basket analysis for a large set of transactions Data mining algorithms (K-means, KNN, and Naive Bayes) Using huge genomic data to sequence DNA and RNA Naive Bayes theorem and Markov chains for data and market prediction Recommendation algorithms and pairwise document similarity Linear regression, Cox regression, and Pearson correlation Allelic frequency and mining DNA Social network analysis (recommendation systems, counting triangles, sentiment analysis)

Hadoop in Practice - Alex Holmes 2014-09-29 Summary Hadoop in Practice, Second Edition provides over 100 tested, instantly useful techniques that will help you conquer big data, using Hadoop. This revised new edition covers changes and new features in the Hadoop core architecture, including MapReduce 2. Brand new chapters cover YARN and integrating Kafka, Impala, and

Spark SQL with Hadoop. You'll also get new and updated techniques for Flume, Sqoop, and Mahout, all of which have seen major new versions recently. In short, this is the most practical, up-to-date coverage of Hadoop available anywhere. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Book It's always a good time to upgrade your Hadoop skills! Hadoop in Practice, Second Edition provides a collection of 104 tested, instantly useful techniques for analyzing real-time streams, moving data securely, machine learning, managing large-scale clusters, and taming big data using Hadoop. This completely revised edition covers changes and new features in Hadoop core, including MapReduce 2 and YARN. You'll pick up hands-on best practices for integrating Spark, Kafka, and Impala with Hadoop, and get new and updated techniques for the latest versions of Flume, Sqoop, and Mahout. In short, this is the most practical, up-to-date coverage of Hadoop available. Readers need to know a programming language like Java and have basic familiarity with Hadoop. What's Inside Thoroughly updated for Hadoop 2 How to write YARN applications Integrate real-time technologies like Storm, Impala, and Spark Predictive analytics using Mahout and RR Readers need to know a programming language like Java and have basic familiarity with Hadoop. About the Author Alex Holmes works on tough big-data problems. He is a software engineer, author, speaker, and blogger specializing in large-scale Hadoop projects. Table of Contents PART 1 BACKGROUND AND FUNDAMENTALS Hadoop in a heartbeat Introduction to YARN PART 2 DATA LOGISTICS Data serialization—working with text and beyond Organizing and optimizing data in HDFS Moving data into and out of Hadoop PART 3 BIG DATA PATTERNS Applying

MapReduce patterns to big data Utilizing data structures and algorithms at scale Tuning, debugging, and testing PART 4 BEYOND MAPREDUCE SQL on Hadoop Writing a YARN application

PHP and MySQL Web Development - Luke Welling 2008-10-01 PHP and MySQL Web Development, Fourth Edition The definitive guide to building database-drive Web applications with PHP and MySQL and MySQL are popular open-source technologies that are ideal for quickly developing database-driven Web applications. PHP is a powerful scripting language designed to enable developers to create highly featured Web applications quickly, and MySQL is a fast, reliable database that integrates well with PHP and is suited for dynamic Internet-based applications. PHP and MySQL Web Development shows how to use these tools together to produce effective, interactive Web applications. It clearly describes the basics of the PHP language, explains how to set up and work with a MySQL database, and then shows how to use PHP to interact with the database and the server. The fourth edition of PHP and MySQL Web Development has been thoroughly updated, revised, and expanded to cover developments in PHP 5 through version 5.3, such as namespaces and closures, as well as features introduced in MySQL 5.1. This is the eBook version of the title. To gain access to the contents on the CD bundled with the printed book, please register your product at informit.com/register

Learn Azure in a Month of Lunches, Second Edition - Iain Foulds 2020-10-06 Learn Azure in a Month of Lunches, Second Edition, is a tutorial on writing, deploying, and running applications in Azure. In it, you'll work through 21 short lessons that give you real-world experience. Each lesson includes a hands-on lab so you can try out and lock in your new skills.

Summary You can be incredibly productive with Azure without mastering every feature, function, and service. Learn Azure in a Month of Lunches, Second Edition gets you up and running quickly, teaching you the most important concepts and tasks in 21 practical bite-sized lessons. As you explore the examples, exercises, and labs, you'll pick up valuable skills immediately and take your first steps to Azure mastery! This fully revised new edition covers core changes to the Azure UI, new Azure features, Azure containers, and the upgraded Azure Kubernetes Service. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Microsoft Azure is vast and powerful, offering virtual servers, application templates, and prebuilt services for everything from data storage to AI. To navigate it all, you need a trustworthy guide. In this book, Microsoft engineer and Azure trainer Iain Foulds focuses on core skills for creating cloud-based applications. About the book Learn Azure in a Month of Lunches, Second Edition, is a tutorial on writing, deploying, and running applications in Azure. In it, you'll work through 21 short lessons that give you real-world experience. Each lesson includes a hands-on lab so you can try out and lock in your new skills. What's inside Understanding Azure beyond point-and-click Securing applications and data Automating your environment Azure services for machine learning, containers, and more About the reader This book is for readers who can write and deploy simple web or client/server applications. About the author Iain Foulds is an engineer and senior content developer with Microsoft. Table of Contents PART 1 - AZURE CORE SERVICES 1 Before you begin 2 Creating a virtual machine 3 Azure Web Apps 4 Introduction to Azure Storage 5 Azure Networking basics PART 2 - HIGH AVAILABILITY AND SCALE 6 Azure Resource Manager 7

High availability and redundancy 8 Load-balancing applications 9 Applications that scale 10 Global databases with Cosmos DB 11 Managing network traffic and routing 12 Monitoring and troubleshooting PART 3 - SECURE BY DEFAULT 13 Backup, recovery, and replication 14 Data encryption 15 Securing information with Azure Key Vault 16 Azure Security Center and updates PART 4 - THE COOL STUFF 17 Machine learning and artificial intelligence 18 Azure Automation 19 Azure containers 20 Azure and the Internet of Things 21 Serverless computing **Practical SQL** - Anthony DeBarros 2018-05-01 Practical SQL is an approachable and fast-paced guide to SQL (Structured Query Language), the standard programming language for defining, organizing, and exploring data in relational databases. The book focuses on using SQL to find the story your data tells, with the popular open-source database PostgreSQL and the pgAdmin interface as its primary tools. You'll first cover the fundamentals of databases and the SQL language, then build skills by analyzing data from the U.S. Census and other federal and state government agencies. With exercises and real-world examples in each chapter, this book will teach even those who have never programmed before all the tools necessary to build powerful databases and access information quickly and efficiently. You'll learn how to: - Create databases and related tables using your own data - Define the right data types for your information - Aggregate, sort, and filter data to find patterns - Use basic math and advanced statistical functions - Identify errors in data and clean them up - Import and export data using delimited text files - Write queries for geographic information systems (GIS) - Create advanced queries and automate tasks Learning SQL doesn't have to be dry and complicated. Practical SQL

delivers clear examples with an easy-to-follow approach to teach you the tools you need to build and manage your own databases. This book uses PostgreSQL, but the SQL syntax is applicable to many database applications, including Microsoft SQL Server and MySQL.

Unreal Engine 4 Game Development in 24 Hours, Sams Teach Yourself - Aram Cookson 2016-06-08
In just 24 lessons of one hour or less, learn how to start using Unreal Engine 4 to build amazing games for Windows, Mac, PS4, Xbox One, iOS, Android, the web, Linux-or all of them! Sams Teach Yourself Unreal Engine 4 Game Development in 24 Hours' straightforward, step-by-step approach shows you how to work with Unreal Engine 4's interface, its workflows, and its most powerful editors and tools. In just hours you'll be creating effects, scripting warfare, implementing physics-even developing for mobile devices and HUDs. Every lesson builds on what you've already learned, giving you a rock-solid foundation for real-world success. Organize new projects and work with the Gameplay Framework Master Unreal's units and control systems Import 3D models and work with the Static Mesh Editor Create new landscapes and use Unreal's foliage system Bring characters and creatures to life with the Persona Editor Apply materials and build lighting Integrate and modify audio with the Unreal Sound Cue Editor Craft particle effects and simulate physics Set up and react to player inputs Build levels and entirely new worlds Get started with powerful Blueprint visual scripting system Script an arcade game from start to finish Create events that respond to player actions Spawn Actors during gameplay Design and create action-based encounters Optimize games for mobile devices and touch-based inputs Build menus with Unreal's UMG UI Designer Prepare your game for deployment Step-by-step instructions carefully walk you

through the most common Unreal Engine 4 game development tasks. Practical, hands-on examples show you how to apply what you learn. Quizzes and Exercises help you test your knowledge and stretch your skills. Notes and tips point out shortcuts and solutions. All the project files and assets you'll need are available for download, including "before-and-after" files demonstrating initial setup and proper completion for every exercise.

Taming The Big Data Tidal Wave - Bill Franks 2012-04-24 You receive an e-mail. It contains an offer for a complete personal computer system. It seems like the retailer read your mind since you were exploring computers on their web site just a few hours prior.... As you drive to the store to buy the computer bundle, you get an offer for a discounted coffee from the coffee shop you are getting ready to drive past. It says that since you're in the area, you can get 10% off if you stop by in the next 20 minutes.... As you drink your coffee, you receive an apology from the manufacturer of a product that you complained about yesterday on your Facebook page, as well as on the company's web site.... Finally, once you get back home, you receive notice of a special armor upgrade available for purchase in your favorite online video game. It is just what is needed to get past some spots you've been struggling with.... Sound crazy? Are these things that can only happen in the distant future? No. All of these scenarios are possible today! Big data. Advanced analytics. Big data analytics. It seems you can't escape such terms today. Everywhere you turn people are discussing, writing about, and promoting big data and advanced analytics. Well, you can now add this book to the discussion. What is real and what is hype? Such attention can lead one to the suspicion that perhaps the analysis of big data is something that is more hype than substance. While there has been a lot of hype

over the past few years, the reality is that we are in a transformative era in terms of analytic capabilities and the leveraging of massive amounts of data. If you take the time to cut through the sometimes-over-zealous hype present in the media, you'll find something very real and very powerful underneath it. With big data, the hype is driven by genuine excitement and anticipation of the business and consumer benefits that analyzing it will yield over time. Big data is the next wave of new data sources that will drive the next wave of analytic innovation in business, government, and academia. These innovations have the potential to radically change how organizations view their business. The analysis that big data enables will lead to decisions that are more informed and, in some cases, different from what they are today. It will yield insights that many can only dream about today. As you'll see, there are many consistencies with the requirements to tame big data and what has always been needed to tame new data sources. However, the additional scale of big data necessitates utilizing the newest tools, technologies, methods, and processes. The old way of approaching analysis just won't work. It is time to evolve the world of advanced analytics to the next level. That's what this book is about. Taming the Big Data Tidal Wave isn't just the title of this book, but rather an activity that will determine which businesses win and which lose in the next decade. By preparing and taking the initiative, organizations can ride the big data tidal wave to success rather than being pummeled underneath the crushing surf. What do you need to know and how do you prepare in order to start taming big data and generating exciting new analytics from it? Sit back, get comfortable, and prepare to find out!

Oracle PL/SQL Programming - Steven Feuerstein 2002 The authors have revised and updated this bestseller to

include both the Oracle8i and new Oracle9i Internet-savvy database products.

Programming Pig - Alan Gates 2016-11-09
For many organizations, Hadoop is the first step for dealing with massive amounts of data. The next step? Processing and analyzing datasets with the Apache Pig scripting platform. With Pig, you can batch-process data without having to create a full-fledged application, making it easy to experiment with new datasets. Updated with use cases and programming examples, this second edition is the ideal learning tool for new and experienced users alike. You'll find comprehensive coverage on key features such as the Pig Latin scripting language and the Grunt shell. When you need to analyze terabytes of data, this book shows you how to do it efficiently with Pig. Delve into Pig's data model, including scalar and complex data types Write Pig Latin scripts to sort, group, join, project, and filter your data Use Grunt to work with the Hadoop Distributed File System (HDFS) Build complex data processing pipelines with Pig's macros and modularity features Embed Pig Latin in Python for iterative processing and other advanced tasks Use Pig with Apache Tez to build high-performance batch and interactive data processing applications Create your own load and store functions to handle data formats and storage mechanisms

Spark Cookbook - Rishi Yadav 2015-07-27
By introducing in-memory persistent storage, Apache Spark eliminates the need to store intermediate data in filesystems, thereby increasing processing speed by up to 100 times. This book will focus on how to analyze large and complex sets of data. Starting with installing and configuring Apache Spark with various cluster managers, you will cover setting up development environments. You will then cover various recipes to perform interactive queries using Spark SQL and real-time

streaming with various sources such as Twitter Stream and Apache Kafka. You will then focus on machine learning, including supervised learning, unsupervised learning, and recommendation engine algorithms. After mastering graph processing using GraphX, you will cover various recipes for cluster optimization and troubleshooting. *Learning Spark SQL* - Aurobindo Sarkar 2017-09-07 Design, implement, and deliver successful streaming applications, machine learning pipelines and graph applications using Spark SQL API About This Book Learn about the design and implementation of streaming applications, machine learning pipelines, deep learning, and large-scale graph processing applications using Spark SQL APIs and Scala. Learn data exploration, data munging, and how to process structured and semi-structured data using real-world datasets and gain hands-on exposure to the issues and challenges of working with noisy and "dirty" real-world data. Understand design considerations for scalability and performance in web-scale Spark application architectures. Who This Book Is For If you are a developer, engineer, or an architect and want to learn how to use Apache Spark in a web-scale project, then this is the book for you. It is assumed that you have prior knowledge of SQL querying. A basic programming knowledge with Scala, Java, R, or Python is all you need to get started with this book. What You Will Learn Familiarize yourself with Spark SQL programming, including working with DataFrame/Dataset API and SQL Perform a series of hands-on exercises with different types of data sources, including CSV, JSON, Avro, MySQL, and MongoDB Perform data quality checks, data visualization, and basic statistical analysis tasks Perform data munging tasks on publically available datasets Learn how to use Spark SQL and Apache Kafka to build streaming applications Learn key performance-tuning tips

and tricks in Spark SQL applications Learn key architectural components and patterns in large-scale Spark SQL applications In Detail In the past year, Apache Spark has been increasingly adopted for the development of distributed applications. Spark SQL APIs provide an optimized interface that helps developers build such applications quickly and easily. However, designing web-scale production applications using Spark SQL APIs can be a complex task. Hence, understanding the design and implementation best practices before you start your project will help you avoid these problems. This book gives an insight into the engineering practices used to design and build real-world, Spark-based applications. The book's hands-on examples will give you the required confidence to work on any future projects you encounter in Spark SQL. It starts by familiarizing you with data exploration and data munging tasks using Spark SQL and Scala. Extensive code examples will help you understand the methods used to implement typical use-cases for various types of applications. You will get a walkthrough of the key concepts and terms that are common to streaming, machine learning, and graph applications. You will also learn key performance-tuning details including Cost Based Optimization (Spark 2.2) in Spark SQL applications. Finally, you will move on to learning how such systems are architected and deployed for a successful delivery of your project. Style and approach This book is a hands-on guide to designing, building, and deploying Spark SQL-centric production applications at scale.

Data Analytics with Spark Using Python - Jeffrey Aven 2018-06-18
Solve Data Analytics Problems with Spark, PySpark, and Related Open Source Tools Spark is at the heart of today's Big Data revolution, helping data professionals supercharge efficiency and performance in a wide range of data processing

and analytics tasks. In this guide, Big Data expert Jeffrey Aven covers all you need to know to leverage Spark, together with its extensions, subprojects, and wider ecosystem. Aven combines a language-agnostic introduction to foundational Spark concepts with extensive programming examples utilizing the popular and intuitive PySpark development environment. This guide's focus on Python makes it widely accessible to large audiences of data professionals, analysts, and developers—even those with little Hadoop or Spark experience. Aven's broad coverage ranges from basic to advanced Spark programming, and Spark SQL to machine learning. You'll learn how to efficiently manage all forms of data with Spark: streaming, structured, semi-structured, and unstructured. Throughout, concise topic overviews quickly get you up to speed, and extensive hands-on exercises prepare you to solve real problems. Coverage includes: • Understand Spark's evolving role in the Big Data and Hadoop ecosystems • Create Spark clusters using various deployment modes • Control and optimize the operation of Spark clusters and applications • Master Spark Core RDD API programming techniques • Extend, accelerate, and optimize Spark routines with advanced API platform constructs, including shared variables, RDD storage, and partitioning • Efficiently integrate Spark with both SQL and nonrelational data stores • Perform stream processing and messaging with Spark Streaming and Apache Kafka • Implement predictive modeling with SparkR and Spark MLlib *SQL in 10 Minutes a Day, Sams Teach Yourself* - Ben Forta 2019-10-29 Sams Teach Yourself SQL in 10 Minutes offers straightforward, practical answers when you need fast results. By working through the book's 22 lessons of 10 minutes or less, you'll learn what you need to know to take advantage of the SQL language. Lessons cover IBM DB2, Microsoft SQL Server and

SQL Server Express, MariaDB, MySQL, Oracle and Oracle express, PostgreSQL, and SQLite. Full-color code examples help you understand how SQL statements are structured Tips point out shortcuts and solutions Cautions help you avoid common pitfalls Notes explain additional concepts, and provide additional information 10 minutes is all you need to learn how to... Use the major SQL statements Construct complex SQL statements using multiple clauses and operators Retrieve, sort, and format database contents Pinpoint the data you need using a variety of filtering techniques Use aggregate functions to summarize data Join two or more related tables Insert, update, and delete data Create and alter database tables Work with views, stored procedures, and more
Spark: The Definitive Guide - Bill Chambers 2018-02-08 Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You'll explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and Datasets—Spark's core APIs—through worked examples Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark

clusters and applications Learn the power of Structured Streaming, Spark's stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation

Hadoop 2 Quick-Start Guide - Douglas Eadline 2015-10-28 Get Started Fast with Apache Hadoop® 2, YARN, and Today's Hadoop Ecosystem With Hadoop 2.x and YARN, Hadoop moves beyond MapReduce to become practical for virtually any type of data processing. Hadoop 2.x and the Data Lake concept represent a radical shift away from conventional approaches to data usage and storage. Hadoop 2.x installations offer unmatched scalability and breakthrough extensibility that supports new and existing Big Data analytics processing methods and models. Hadoop® 2 Quick-Start Guide is the first easy, accessible guide to Apache Hadoop 2.x, YARN, and the modern Hadoop ecosystem. Building on his unsurpassed experience teaching Hadoop and Big Data, author Douglas Eadline covers all the basics you need to know to install and use Hadoop 2 on personal computers or servers, and to navigate the powerful technologies that complement it. Eadline concisely introduces and explains every key Hadoop 2 concept, tool, and service, illustrating each with a simple "beginning-to-end" example and identifying trustworthy, up-to-date resources for learning more. This guide is ideal if you want to learn about Hadoop 2 without getting mired in technical details. Douglas Eadline will bring you up to speed quickly, whether you're a user, admin, devops specialist, programmer, architect, analyst, or data scientist. Coverage Includes Understanding what Hadoop 2 and YARN do, and how they improve on Hadoop 1 with MapReduce Understanding Hadoop-based Data Lakes versus RDBMS Data Warehouses Installing Hadoop 2 and core services on Linux machines, virtualized sandboxes, or clusters

Exploring the Hadoop Distributed File System (HDFS) Understanding the essentials of MapReduce and YARN application programming Simplifying programming and data movement with Apache Pig, Hive, Sqoop, Flume, Oozie, and HBase Observing application progress, controlling jobs, and managing workflows Managing Hadoop efficiently with Apache Ambari–including recipes for HDFS to NFSv3 gateway, HDFS snapshots, and YARN configuration Learning basic Hadoop 2 troubleshooting, and installing Apache Hue and Apache Spark

Sams Teach Yourself SQL in 10 Minutes - Ben Forta 2004 Explains how to use Structured Query Language to work within a relational database system, including information retrieval, security, data manipulation, and user management.

Business Analytics - Thomas W. Jackson 2018-09-21 This innovative new textbook, co-authored by an established academic and a leading practitioner, is the first to bring together issues of cloud computing, business intelligence and big data analytics in order to explore how organisations use cloud technology to analyse data and make decisions. In addition to offering an up-to-date exploration of key issues relating to data privacy and ethics, information governance, and the future of analytics, the text describes the options available in deploying analytic solutions to the cloud and draws on real-world, international examples from companies such as Rolls Royce, Lego, Volkswagen and Samsung. Combining academic and practitioner perspectives that are crucial to the understanding of this growing field, Business Analytics acts an ideal core text for undergraduate, postgraduate and MBA modules on Big Data, Business and Data Analytics, and Business Intelligence, as well as functioning as a supplementary text for modules in Marketing Analytics. The book is also an invaluable resource for

practitioners and will quickly enable the next generation of 'Information Builders' within organisations to understand innovative cloud based-analytic solutions.

JSON at Work - Tom Marrs 2017-06-19
JSON is becoming the backbone for meaningful data interchange over the internet. This format is now supported by an entire ecosystem of standards, tools, and technologies for building truly elegant, useful, and efficient applications. With this hands-on guide, author and architect Tom Marrs shows you how to build enterprise-class applications and services by leveraging JSON tooling and message/document design. JSON at Work provides application architects and developers with guidelines, best practices, and use cases, along with lots of real-world examples and code samples. You'll start with a comprehensive JSON overview, explore the JSON ecosystem, and then dive into JSON's use in the enterprise. Get acquainted with JSON basics and learn how to model JSON data Learn how to use JSON with Node.js, Ruby on Rails, and Java Structure JSON documents with JSON Schema to design and test APIs Search the contents of JSON documents with JSON Search tools Convert JSON documents to other data formats with JSON Transform tools Compare JSON-based hypermedia formats, including HAL and jsonapi Leverage MongoDB to store and access JSON documents Use Apache Kafka to exchange JSON-based messages between services

Deep Learning - Josh Patterson 2017-07-28
Although interest in machine learning has reached a high point, lofty expectations often scuttle projects before they get very far. How can machine learning—especially deep neural networks—make a real difference in your organization? This hands-on guide not only provides the most practical information available on the subject, but also helps you get started

building efficient deep learning networks. Authors Adam Gibson and Josh Patterson provide theory on deep learning before introducing their open-source Deeplearning4j (DL4J) library for developing production-class workflows. Through real-world examples, you'll learn methods and strategies for training deep network architectures and running deep learning workflows on Spark and Hadoop with DL4J. Dive into machine learning concepts in general, as well as deep learning in particular Understand how deep networks evolved from neural network fundamentals Explore the major deep network architectures, including Convolutional and Recurrent Learn how to map specific deep networks to the right problem Walk through the fundamentals of tuning general neural networks and specific deep network architectures Use vectorization techniques for different data types with DataVec, DL4J's workflow tool Learn how to use DL4J natively on Spark and Hadoop

Interactive Data Visualization for the Web - Scott Murray 2013-03-11 Author Scott Murray teaches you the fundamental concepts and methods of D3, a JavaScript library that lets you express data visually in a web browser

Sams Teach Yourself SQL in 24 Hours - Ryan Stephens 2008-05-30 In just 24 lessons of one hour or less, you will learn professional techniques to design and build efficient databases and query them to extract useful information. Using a straightforward, step-by-step approach, each lesson builds on the previous one, allowing you to learn the essentials of ANSI SQL from the ground up. Example code demonstrates the authors' professional techniques, while exercises written for MySQL offer the reader hands-on learning with an open-source database. Included are advanced techniques for using views, managing transactions, database administration, and

extending SQL. Step-by-step instructions carefully walk you through the most common SQL tasks. Q&As, Quizzes, and Exercises at the end of each chapter help you test your knowledge. Notes and Tips point out shortcuts and solutions. New terms are clearly defined and explained. Learn how to… Use SQL-2003, the latest standard for the Structured Query Language Design and deploy efficient, secure databases Build advanced queries for information retrieval Sort, group, and summarize information for best presentation Tune databases and queries for maximum performance Understand database administration and security techniques For more than ten years the authors have studied, applied, and documented the SQL standard and its application to critical database systems. Ryan Stephens and Ron Plew are entrepreneurs, speakers, and cofounders of Perpetual Technologies, Inc. (PTI), a fast-growing IT management and consulting firm which specializes in database technologies. They taught database courses for Indiana University–Purdue University in Indianapolis for five years and have authored more than a dozen books on Oracle, SQL, database design, and the high availability of critical systems. Arie D. Jones is Senior SQL Server database administrator and analyst for PTI. He is a regular speaker at technical events and has authored several books and articles. Category: Database Covers: ANSI SQL User Level: Beginning–Intermediate Register your book at informit.com/title/9780672330 186 for convenient access to updates and corrections as they become available.

Hadoop in 24 Hours, Sams Teach Yourself - Jeffrey Aven 2017-04-07
Apache Hadoop is the technology at the heart of the Big Data revolution, and Hadoop skills are in enormous demand. Now, in just 24 lessons of one hour or less, you can learn all the skills and

techniques you'll need to deploy each key component of a Hadoop platform in your local environment or in the cloud, building a fully functional Hadoop cluster and using it with real programs and datasets. Each short, easy lesson builds on all that's come before, helping you master all of Hadoop's essentials, and extend it to meet your unique challenges. Apache Hadoop in 24 Hours, Sams Teach Yourself covers all this, and much more: Understanding Hadoop and the Hadoop Distributed File System (HDFS) Importing data into Hadoop, and process it there Mastering basic MapReduce Java programming, and using advanced MapReduce API concepts Making the most of Apache Pig and Apache Hive Implementing and administering YARN Taking advantage of the full Hadoop ecosystem Managing Hadoop clusters with Apache Ambari Working with the Hadoop User Environment (HUE) Scaling, securing, and troubleshooting Hadoop environments Integrating Hadoop into the enterprise Deploying Hadoop in the cloud Getting started with Apache Spark Step-by-step instructions walk you through common questions, issues, and tasks; Q-and-As, Quizzes, and Exercises build and test your knowledge; "Did You Know?" tips offer insider advice and shortcuts; and "Watch Out!" alerts help you avoid pitfalls. By the time you're finished, you'll be comfortable using Apache Hadoop to solve a wide spectrum of Big Data problems. **Armored** - Mark Greaney 2022-07-05 A novel inspired by #1 New York Times best-selling author Mark Greaney's Audible Original drama, Armored. Joshua Duffy is a Close Protection Agent—a professional bodyguard—and he's one of the world's elite operatives. That is, he was until his last mission in Lebanon. Against all odds, Josh got his primary out alive, but the cost was high. Josh lost his lower left leg. There's not much call for an elite bodyguard with such an injury.

So, Josh has to support his family working as a mall cop in New Jersey. For a man like Josh, this is purgatory on earth, but miracles can occur even in Paramus. A lucky run-in with an old comrade promises to get Josh back in the field for one last job. The UN is sending a peace mission into the Sierra Madre mountains in Mexico, an area so dangerous it's known as Espinazo del Diablo—the Devil's Spine. Only a fool would think they could broker peace between the homicidal drug cartels in the region, and only a madman would sign on to keep those fools alive.

Big Data Analytics with Microsoft HDInsight in 24 Hours, Sams Teach Yourself
- Manpreet Singh 2015-11-12 Sams Teach Yourself Big Data Analytics with Microsoft HDInsight in 24 Hours In just 24 lessons of one hour or less, Sams Teach Yourself Big Data Analytics with Microsoft HDInsight in 24 Hours helps you leverage Hadoop's power on a flexible, scalable cloud platform using Microsoft's newest business intelligence, visualization, and productivity tools. This book's straightforward, step-by-step approach shows you how to provision, configure, monitor, and troubleshoot HDInsight and use Hadoop cloud services to solve real analytics problems. You'll gain more of Hadoop's benefits, with less complexity–even if you're completely new to Big Data analytics. Every lesson builds on what you've already learned, giving you a rock-solid foundation for real-world success. Practical, hands-on examples show you how to apply what you learn Quizzes and exercises help you test your knowledge and stretch your skills Notes and tips point out shortcuts and solutions Learn how to… · Master core Big Data and NoSQL concepts, value propositions, and use cases · Work with key Hadoop features, such as HDFS2 and YARN · Quickly install, configure, and monitor Hadoop (HDInsight) clusters in the cloud · Automate provisioning, customize clusters, install additional Hadoop projects,

and administer clusters · Integrate, analyze, and report with Microsoft BI and Power BI · Automate workflows for data transformation, integration, and other tasks · Use Apache HBase on HDInsight · Use Sqoop or SSIS to move data to or from HDInsight · Perform R-based statistical computing on HDInsight datasets · Accelerate analytics with Apache Spark · Run real-time analytics on high-velocity data streams · Write MapReduce, Hive, and Pig programs Register your book at informit.com/register for convenient access to downloads, updates, and corrections as they become available.

Apache Spark in 24 Hours - Andrea Avena 2017

Learning Spark - Holden Karau 2015-01-28 Data in all domains is getting bigger. How can you work with it efficiently? Recently updated for Spark 1.3, this book introduces Apache Spark, the open source cluster computing system that makes data analytics fast to write and fast to run. With Spark, you can tackle big datasets quickly through simple APIs in Python, Java, and Scala. This edition includes new information on Spark SQL, Spark Streaming, setup, and Maven coordinates. Written by the developers of Spark, this book will have data scientists and engineers up and running in no time. You'll learn how to express parallel jobs with just a few lines of code, and cover applications from simple batch jobs to stream processing and machine learning. Quickly dive into Spark capabilities such as distributed datasets, in-memory caching, and the interactive shell Leverage Spark's powerful built-in libraries, including Spark SQL, Spark Streaming, and MLlib Use one programming paradigm instead of mixing and matching tools like Hive, Hadoop, Mahout, and Storm Learn how to deploy interactive, batch, and streaming applications Connect to data sources including HDFS, Hive, JSON, and S3 Master advanced topics like

data partitioning and shared variables

Apache Spark in 24 Hours, Sams Teach Yourself - Jeffrey Aven 2016-08-31 Apache Spark is a fast, scalable, and flexible open source distributed processing engine for big data systems and is one of the most active open source big data projects to date. In just 24 lessons of one hour or less, Sams Teach Yourself Apache Spark in 24 Hours helps you build practical Big Data solutions that leverage Spark's amazing speed, scalability, simplicity, and versatility. This book's straightforward, step-by-step approach shows you how to deploy, program, optimize, manage, integrate, and extend Spark–now, and for years to come. You'll discover how to create powerful solutions encompassing cloud computing, real-time stream processing, machine learning, and more. Every lesson builds on what you've already learned, giving you a rock-solid foundation for real-world success. Whether you are a data analyst, data engineer, data scientist, or data steward, learning Spark will help you to advance your career or embark on a new career in the booming area of Big Data. Learn how to • Discover what Apache Spark does and how it fits into the Big Data landscape • Deploy and run Spark locally or in the cloud • Interact with Spark from the shell • Make the most of the Spark Cluster Architecture • Develop Spark applications with Scala and functional Python • Program with the Spark API, including transformations and actions • Apply practical data engineering/analysis approaches designed for Spark • Use Resilient Distributed Datasets (RDDs) for caching, persistence, and output • Optimize Spark solution performance • Use Spark with SQL (via Spark SQL) and with NoSQL (via Cassandra) • Leverage cutting-edge functional programming techniques • Extend Spark with streaming, R, and Sparkling Water • Start building Spark-based machine

learning and graph-processing applications • Explore advanced messaging technologies, including Kafka • Preview and prepare for Spark's next generation of innovations Instructions walk you through common questions, issues, and tasks; Q-and-As, Quizzes, and Exercises build and test your knowledge; "Did You Know?" tips offer insider advice and shortcuts; and "Watch Out!" alerts help you avoid pitfalls. By the time you're finished, you'll be comfortable using Apache Spark to solve a wide spectrum of Big Data problems.
Spark in Action - Marko Bonaci 2016-11-03
Summary Spark in Action teaches you the theory and skills you need to effectively handle batch and streaming data using Spark. Fully updated for Spark 2.0. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Big data systems distribute datasets across clusters of machines, making it a challenge to efficiently query, stream, and interpret them. Spark can help. It is a processing system designed specifically for distributed data. It provides easy-to-use interfaces, along with the performance you need for production-quality analytics and machine learning. Spark 2 also adds improved programming APIs, better performance, and countless other upgrades. About the Book Spark in Action teaches you the theory and skills you need to effectively handle batch and streaming data using Spark. You'll get comfortable with the Spark CLI as you work through a few introductory examples. Then, you'll start programming Spark using its core APIs. Along the way, you'll work with structured data using Spark SQL, process near-real-time streaming data, apply machine learning algorithms, and munge graph data using Spark GraphX. For a zero-effort startup, you can download the preconfigured virtual machine ready for you to try the book's code. What's

Inside Updated for Spark 2.0 Real-life case studies Spark DevOps with Docker Examples in Scala, and online in Java and Python About the Reader Written for experienced programmers with some background in big data or machine learning. About the Authors Petar Zečević and Marko Bonaći are seasoned developers heavily involved in the Spark community. Table of Contents PART 1 - FIRST STEPS Introduction to Apache Spark Spark fundamentals Writing Spark applications The Spark API in depth PART 2 - MEET THE SPARK FAMILY Sparkling queries with Spark SQL Ingesting data with Spark Streaming Getting smart with MLlib ML: classification and clustering Connecting the dots with GraphX PART 3 - SPARK OPS Running Spark Running on a Spark standalone cluster Running on YARN and Mesos PART 4 - BRINGING IT TOGETHER Case study: real-time dashboard Deep learning on Spark with H2O
SQL in 10 Minutes, Sams Teach Yourself - Ben Forta 2012-10-25 Sams Teach Yourself SQL in 10 Minutes, Fourth Edition New full-color code examples help you see how SQL statements are structured Whether you're an application developer, database administrator, web application designer, mobile app developer, or Microsoft Office users, a good working knowledge of SQL is an important part of interacting with databases. And Sams Teach Yourself SQL in 10 Minutes offers the straightforward, practical answers you need to help you do your job. Expert trainer and popular author Ben Forta teaches you just the parts of SQL you need to know–starting with simple data retrieval and quickly going on to more complex topics including the use of joins, subqueries, stored procedures, cursors, triggers, and table constraints. You'll learn methodically, systematically, and simply–in 22 short, quick lessons that will each take only 10 minutes or less to complete. With the Fourth Edition of this

worldwide bestseller, the book has been thoroughly updated, expanded, and improved. Lessons now cover the latest versions of IBM DB2, Microsoft Access, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, SQLite, MariaDB, and Apache Open Office Base. And new full-color SQL code listings help the beginner clearly see the elements and structure of the language. 10 minutes is all you need to learn how to... Use the major SQL statements Construct complex SQL statements using multiple clauses and operators Retrieve, sort, and format database contents Pinpoint the data you need using a variety of filtering techniques Use aggregate functions to summarize data Join two or more related tables Insert, update, and delete data Create and alter database tables Work with views, stored procedures, and more Table of Contents 1 Understanding SQL 2 Retrieving Data 3 Sorting Retrieved Data 4 Filtering Data 5 Advanced Data Filtering 6 Using Wildcard Filtering 7 Creating Calculated Fields 8 Using Data Manipulation Functions 9 Summarizing Data 10 Grouping Data 11 Working with Subqueries 12 Joining Tables 13 Creating Advanced Joins 14 Combining Queries 15 Inserting Data 16 Updating and Deleting Data 17 Creating and Manipulating Tables 18 Using Views 19 Working with Stored Procedures 20 Managing Transaction Processing 21 Using Cursors 22 Understanding Advanced SQL Features Appendix A: Sample Table Scripts Appendix B: Working in Popular Applications Appendix C : SQL Statement Syntax Appendix D: Using SQL Datatypes Appendix E: SQL Reserved Words