

HANDS-ON DATA SCIENCE WITH SQL SERVER 2017

HANDS-ON DATA SCIENCE WITH SQL SERVER 2017 OFFERS A PRACTICAL APPROACH TO LEVERAGING THE POWERFUL CAPABILITIES OF MICROSOFT'S SQL SERVER PLATFORM FOR ADVANCED DATA ANALYTICS AND MACHINE LEARNING. THIS ARTICLE EXPLORES HOW SQL SERVER 2017 INTEGRATES DATA SCIENCE WORKFLOWS DIRECTLY WITHIN THE DATABASE ENVIRONMENT, ENABLING EFFICIENT DATA MANIPULATION, PREDICTIVE MODELING, AND REAL-TIME ANALYTICS. EMPHASIZING THE HANDS-ON ASPECT, THE DISCUSSION INCLUDES KEY FEATURES SUCH AS R AND PYTHON INTEGRATION, IN-DATABASE MACHINE LEARNING, AND SCALABLE DATA PROCESSING. READERS WILL GAIN INSIGHTS INTO SETTING UP A DATA SCIENCE ENVIRONMENT, EXECUTING ANALYTICAL MODELS, AND OPTIMIZING PERFORMANCE FOR LARGE DATASETS. WITH THE GROWING IMPORTANCE OF DATA-DRIVEN DECISION-MAKING, MASTERING HANDS-ON DATA SCIENCE WITH SQL SERVER 2017 IS ESSENTIAL FOR DATA PROFESSIONALS AIMING TO HARNESS ENTERPRISE DATA EFFECTIVELY. THE FOLLOWING SECTIONS OUTLINE THE ESSENTIAL COMPONENTS AND PRACTICAL STEPS TO IMPLEMENT ROBUST DATA SCIENCE SOLUTIONS USING SQL SERVER 2017.

- OVERVIEW OF SQL SERVER 2017 FOR DATA SCIENCE
- INTEGRATING R AND PYTHON FOR ADVANCED ANALYTICS
- IN-DATABASE MACHINE LEARNING CAPABILITIES
- DATA PREPARATION AND TRANSFORMATION TECHNIQUES
- DEPLOYING AND OPERATIONALIZING MODELS
- PERFORMANCE OPTIMIZATION AND SCALABILITY

OVERVIEW OF SQL SERVER 2017 FOR DATA SCIENCE

SQL SERVER 2017 PROVIDES A COMPREHENSIVE PLATFORM FOR DATA SCIENCE BY COMBINING TRADITIONAL DATABASE MANAGEMENT WITH ADVANCED ANALYTICS CAPABILITIES. IT SUPPORTS RUNNING COMPLEX QUERIES, MANAGING LARGE VOLUMES OF DATA, AND INTEGRATING STATISTICAL AND MACHINE LEARNING TOOLS DIRECTLY WITHIN THE DATABASE ENGINE. THIS INTEGRATION REDUCES THE NEED TO MOVE DATA BETWEEN SYSTEMS, THEREBY INCREASING EFFICIENCY AND SECURITY. SQL SERVER 2017 INCLUDES ENHANCED SUPPORT FOR CROSS-PLATFORM DEPLOYMENT, RUNNING ON BOTH WINDOWS AND LINUX, WHICH EXTENDS ITS USABILITY ACROSS DIVERSE ENVIRONMENTS. THE PLATFORM'S COMPATIBILITY WITH INDUSTRY-STANDARD PROGRAMMING LANGUAGES MAKES IT A VERSATILE CHOICE FOR DATA SCIENTISTS AND DATABASE ADMINISTRATORS ALIKE.

KEY FEATURES SUPPORTING DATA SCIENCE

SEVERAL FEATURES MAKE SQL SERVER 2017 PARTICULARLY SUITABLE FOR HANDS-ON DATA SCIENCE PROJECTS:

- **INTEGRATED R AND PYTHON SUPPORT:** ENABLES RUNNING SCRIPTS AND MODELS INSIDE THE DATABASE ENGINE.
- **MACHINE LEARNING SERVICES:** PROVIDES SCALABLE, IN-DATABASE ANALYTICS WITHOUT DATA MOVEMENT.
- **GRAPH DATA PROCESSING:** FACILITATES ANALYSIS OF RELATIONSHIPS WITHIN COMPLEX DATASETS.
- **ADVANCED DATA SECURITY:** ENSURES COMPLIANCE AND PROTECTION OF SENSITIVE DATA.
- **HIGH AVAILABILITY AND SCALABILITY:** SUPPORTS ENTERPRISE-GRADE DEPLOYMENTS AND REAL-TIME ANALYTICS.

INTEGRATING R AND PYTHON FOR ADVANCED ANALYTICS

ONE OF THE STANDOUT CAPABILITIES OF SQL SERVER 2017 IS THE SEAMLESS INTEGRATION OF R AND PYTHON PROGRAMMING LANGUAGES, BOTH WIDELY USED IN THE DATA SCIENCE COMMUNITY. THIS FEATURE ALLOWS DATA SCIENTISTS TO EXECUTE SCRIPTS DIRECTLY WITHIN THE SQL SERVER ENVIRONMENT, COMBINING THE POWER OF THESE LANGUAGES WITH THE ROBUSTNESS OF A RELATIONAL DATABASE. THIS INTEGRATION SUPPORTS THE DEVELOPMENT, TESTING, AND DEPLOYMENT OF MACHINE LEARNING MODELS WITHOUT THE OVERHEAD OF EXPORTING DATA TO EXTERNAL PLATFORMS.

SETTING UP R AND PYTHON IN SQL SERVER

TO UTILIZE R AND PYTHON WITHIN SQL SERVER 2017, THE MACHINE LEARNING SERVICES FEATURE MUST BE INSTALLED AND ENABLED. THIS SETUP INVOLVES CONFIGURING EXTERNAL SCRIPT EXECUTION PERMISSIONS AND ENSURING THE NECESSARY RUNTIME ENVIRONMENTS ARE IN PLACE. ONCE CONFIGURED, USERS CAN INVOKE R OR PYTHON SCRIPTS USING THE `SP_EXECUTE_EXTERNAL_SCRIPT` STORED PROCEDURE, INTEGRATING DATA PROCESSING AND MODEL SCORING IN A STREAMLINED WORKFLOW.

BENEFITS OF IN-DATABASE ANALYTICS WITH R AND PYTHON

EXECUTING ANALYTICS WITHIN SQL SERVER OFFERS SEVERAL ADVANTAGES:

- **REDUCED DATA MOVEMENT:** ELIMINATES THE NEED TO TRANSFER LARGE DATASETS TO EXTERNAL TOOLS.
- **IMPROVED SECURITY:** DATA REMAINS PROTECTED WITHIN THE DATABASE BOUNDARIES.
- **PERFORMANCE GAINS:** LEVERAGES SQL SERVER'S PARALLEL PROCESSING AND INDEXING.
- **OPERATIONAL SIMPLICITY:** CONSOLIDATES DATA MANAGEMENT AND ANALYTICS IN ONE PLATFORM.

IN-DATABASE MACHINE LEARNING CAPABILITIES

SQL SERVER 2017'S IN-DATABASE MACHINE LEARNING SERVICES ALLOW USERS TO BUILD, TRAIN, AND DEPLOY PREDICTIVE MODELS WITHOUT LEAVING THE DATABASE ENVIRONMENT. THIS APPROACH SUPPORTS A VARIETY OF ALGORITHMS AND STATISTICAL METHODS ACCESSIBLE THROUGH R AND PYTHON, INCLUDING CLASSIFICATION, REGRESSION, CLUSTERING, AND TIME SERIES ANALYSIS. BY BRINGING MACHINE LEARNING CLOSER TO THE DATA, SQL SERVER ENHANCES BOTH THE SPEED AND SCALABILITY OF ANALYTIC WORKFLOWS.

SUPPORTED ALGORITHMS AND MODELS

THE PLATFORM SUPPORTS DIVERSE MACHINE LEARNING TECHNIQUES THAT ARE ESSENTIAL FOR HANDS-ON DATA SCIENCE PROJECTS. USERS CAN IMPLEMENT ALGORITHMS SUCH AS DECISION TREES, RANDOM FORESTS, LOGISTIC REGRESSION, NEURAL NETWORKS, AND SUPPORT VECTOR MACHINES DIRECTLY WITHIN SQL SERVER. THESE CAPABILITIES ENABLE ORGANIZATIONS TO CONDUCT PREDICTIVE ANALYTICS, ANOMALY DETECTION, AND CUSTOMER SEGMENTATION EFFICIENTLY.

MODEL TRAINING AND SCORING INSIDE SQL SERVER

TRAINING MODELS WITHIN SQL SERVER INVOLVES USING STORED PROCEDURES THAT INVOKE R OR PYTHON SCRIPTS TO PROCESS DATA AND GENERATE MODEL OBJECTS. THESE MODELS CAN THEN BE STORED IN THE DATABASE FOR LATER USE. SCORING, OR APPLYING MODELS TO NEW DATA, IS PERFORMED IN REAL-TIME THROUGH SIMILAR PROCEDURES, SUPPORTING OPERATIONAL ANALYTICS AND DECISION-MAKING PROCESSES SEAMLESSLY INTEGRATED INTO BUSINESS APPLICATIONS.

DATA PREPARATION AND TRANSFORMATION TECHNIQUES

EFFECTIVE DATA SCIENCE BEGINS WITH THOROUGH DATA PREPARATION, AND SQL SERVER 2017 OFFERS ROBUST TOOLS FOR CLEANSING, TRANSFORMING, AND ENRICHING RAW DATA. UTILIZING T-SQL ALONGSIDE R AND PYTHON SCRIPTS, DATA PROFESSIONALS CAN PERFORM COMPLEX TRANSFORMATIONS, FEATURE ENGINEERING, AND AGGREGATION TASKS DIRECTLY WITHIN THE DATABASE. THIS INTEGRATED APPROACH SIMPLIFIES WORKFLOW AND REDUCES LATENCY IN DATA PROCESSING PIPELINES.

COMMON DATA TRANSFORMATION TASKS

TYPICAL DATA PREPARATION ACTIVITIES SUPPORTED BY SQL SERVER 2017 INCLUDE:

- DATA CLEANING AND VALIDATION TO ENSURE QUALITY AND CONSISTENCY.
- NORMALIZATION AND SCALING OF FEATURES FOR MACHINE LEARNING MODELS.
- HANDLING MISSING VALUES THROUGH IMPUTATION TECHNIQUES.
- GENERATING NEW FEATURES USING STATISTICAL OR DOMAIN-SPECIFIC CALCULATIONS.
- AGGREGATING AND SUMMARIZING DATA FOR TREND ANALYSIS.

LEVERAGING SQL SERVER TOOLS FOR ETL PROCESSES

SQL SERVER INTEGRATION SERVICES (SSIS) COMPLEMENTS THE HANDS-ON DATA SCIENCE APPROACH BY ENABLING EFFICIENT EXTRACT, TRANSFORM, LOAD (ETL) OPERATIONS. SSIS PACKAGES CAN AUTOMATE DATA INGESTION AND TRANSFORMATION WORKFLOWS, FEEDING PREPARED DATA INTO SQL SERVER DATABASES FOR FURTHER ANALYSIS AND MODELING. THIS INTEGRATION STREAMLINES DATA PIPELINE MANAGEMENT WITHIN ENTERPRISE ENVIRONMENTS.

DEPLOYING AND OPERATIONALIZING MODELS

DEPLOYING MACHINE LEARNING MODELS IN PRODUCTION IS A CRITICAL STEP IN REALIZING THE VALUE OF DATA SCIENCE INITIATIVES. SQL SERVER 2017 FACILITATES THIS PROCESS BY ALLOWING MODELS DEVELOPED IN R OR PYTHON TO BE EMBEDDED WITHIN STORED PROCEDURES AND FUNCTIONS. THIS CAPABILITY SUPPORTS REAL-TIME SCORING AND INTEGRATION WITH APPLICATIONS, ENABLING DYNAMIC, DATA-DRIVEN DECISION MAKING.

MODEL MANAGEMENT AND VERSIONING

MANAGING MULTIPLE VERSIONS OF MODELS IS ESSENTIAL FOR MAINTAINING ACCURACY AND RELEVANCE. SQL SERVER SUPPORTS STORING SERIALIZED MODELS AS BINARY OBJECTS WITHIN THE DATABASE, ALLOWING FOR VERSION CONTROL AND ROLLBACK IF NEEDED. THIS APPROACH ENSURES MODELS REMAIN CONSISTENT ACROSS DIFFERENT ENVIRONMENTS AND USERS.

REAL-TIME SCORING AND INTEGRATION

BY EMBEDDING SCORING PROCEDURES DIRECTLY WITHIN SQL SERVER, ORGANIZATIONS CAN APPLY PREDICTIVE INSIGHTS IN REAL TIME. THIS INTEGRATION SUPPORTS SCENARIOS SUCH AS FRAUD DETECTION, RECOMMENDATION ENGINES, AND CUSTOMER CHURN PREDICTION, WHERE IMMEDIATE RESPONSES ARE CRITICAL. THE ABILITY TO EXECUTE THESE OPERATIONS INSIDE THE DATABASE REDUCES LATENCY AND IMPROVES OVERALL SYSTEM PERFORMANCE.

PERFORMANCE OPTIMIZATION AND SCALABILITY

HANDLING LARGE DATASETS AND COMPLEX ANALYTICS REQUIRES CAREFUL PERFORMANCE TUNING AND SCALABILITY STRATEGIES. SQL SERVER 2017 OFFERS SEVERAL FEATURES TO OPTIMIZE HANDS-ON DATA SCIENCE WORKFLOWS, ENSURING EFFICIENT RESOURCE UTILIZATION AND FAST QUERY EXECUTION. THESE ENHANCEMENTS SUPPORT THE DEMANDS OF ENTERPRISE-GRADE DATA SCIENCE PROJECTS.

INDEXING AND PARTITIONING STRATEGIES

IMPLEMENTING APPROPRIATE INDEXING AND DATA PARTITIONING IMPROVES QUERY PERFORMANCE BY REDUCING DATA ACCESS TIMES. SQL SERVER 2017 SUPPORTS COLUMNSTORE INDEXES, WHICH ARE PARTICULARLY EFFECTIVE FOR ANALYTICAL WORKLOADS, AS WELL AS TABLE PARTITIONING TO DISTRIBUTE DATA ACROSS PHYSICAL STORAGE.

RESOURCE GOVERNANCE AND PARALLEL PROCESSING

SQL SERVER INCLUDES RESOURCE GOVERNANCE FEATURES THAT MANAGE CPU AND MEMORY ALLOCATION, PREVENTING RESOURCE CONTENTION DURING INTENSIVE DATA SCIENCE OPERATIONS. ADDITIONALLY, THE PLATFORM SUPPORTS PARALLEL QUERY EXECUTION, ACCELERATING PROCESSING TIMES FOR COMPLEX ANALYTICAL QUERIES AND MODEL TRAINING TASKS.

SCALING WITH SQL SERVER ON LINUX AND CLOUD

SQL SERVER 2017'S CROSS-PLATFORM SUPPORT EXTENDS TO CLOUD ENVIRONMENTS, ENABLING SCALABLE DEPLOYMENT OPTIONS. ORGANIZATIONS CAN LEVERAGE CLOUD INFRASTRUCTURE FOR ELASTIC RESOURCE ALLOCATION, SCALING DATA SCIENCE WORKLOADS UP OR DOWN AS NEEDED. THIS FLEXIBILITY SUPPORTS EVOLVING BUSINESS REQUIREMENTS AND LARGE-SCALE DATA PROCESSING NEEDS.

FREQUENTLY ASKED QUESTIONS

WHAT ARE THE KEY FEATURES OF SQL SERVER 2017 FOR DATA SCIENCE?

SQL SERVER 2017 OFFERS INTEGRATED SUPPORT FOR R AND PYTHON, ADVANCED ANALYTICS WITH SQL SERVER MACHINE LEARNING SERVICES, IN-DATABASE COMPUTATIONS, AND SUPPORT FOR GRAPH DATA, MAKING IT A POWERFUL PLATFORM FOR DATA SCIENCE WORKFLOWS.

HOW CAN I PERFORM MACHINE LEARNING TASKS WITHIN SQL SERVER 2017?

YOU CAN USE SQL SERVER MACHINE LEARNING SERVICES TO RUN R AND PYTHON SCRIPTS DIRECTLY INSIDE THE DATABASE. THIS ALLOWS YOU TO BUILD, TRAIN, AND DEPLOY MACHINE LEARNING MODELS WITHOUT MOVING DATA OUTSIDE THE SQL SERVER ENVIRONMENT.

WHAT IS THE BENEFIT OF USING HANDS-ON DATA SCIENCE WITH SQL SERVER 2017?

HANDS-ON DATA SCIENCE WITH SQL SERVER 2017 ENABLES PRACTITIONERS TO LEVERAGE SCALABLE, PRODUCTION-READY ANALYTICS DIRECTLY WITHIN THE DATABASE, IMPROVING PERFORMANCE, SECURITY, AND SIMPLIFYING DEPLOYMENT OF DATA SCIENCE MODELS.

HOW DO I INTEGRATE PYTHON WITH SQL SERVER 2017 FOR DATA SCIENCE

PROJECTS?

SQL SERVER 2017 SUPPORTS PYTHON SCRIPTS THROUGH MACHINE LEARNING SERVICES. YOU CAN WRITE PYTHON CODE INSIDE STORED PROCEDURES OR USE `SP_EXECUTE_EXTERNAL_SCRIPT` TO RUN PYTHON SCRIPTS ON DATA STORED IN SQL SERVER.

CAN SQL SERVER 2017 HANDLE BIG DATA ANALYTICS FOR DATA SCIENCE?

YES, SQL SERVER 2017 SUPPORTS BIG DATA CLUSTERS AND CAN INTEGRATE WITH HADOOP AND SPARK, ALLOWING DATA SCIENTISTS TO PERFORM SCALABLE ANALYTICS AND MANAGE BIG DATA WORKLOADS EFFICIENTLY WITHIN A UNIFIED PLATFORM.

WHAT TOOLS COMPLEMENT SQL SERVER 2017 FOR HANDS-ON DATA SCIENCE?

TOOLS SUCH AS SQL SERVER MANAGEMENT STUDIO (SSMS), AZURE DATA STUDIO, JUPYTER NOTEBOOKS, AND VISUAL STUDIO CODE COMPLEMENT SQL SERVER 2017 BY PROVIDING INTERACTIVE ENVIRONMENTS FOR QUERYING, SCRIPTING, AND DEVELOPING DATA SCIENCE SOLUTIONS.

ADDITIONAL RESOURCES

1. *HANDS-ON DATA SCIENCE WITH SQL SERVER 2017*

THIS BOOK PROVIDES A PRACTICAL APPROACH TO PERFORMING DATA SCIENCE TASKS USING SQL SERVER 2017. IT COVERS DATA PREPARATION, EXPLORATION, MODELING, AND DEPLOYMENT DIRECTLY WITHIN SQL SERVER. READERS WILL LEARN TO LEVERAGE BUILT-IN MACHINE LEARNING SERVICES AND R INTEGRATION TO BUILD POWERFUL ANALYTICS SOLUTIONS. THE BOOK EMPHASIZES REAL-WORLD EXAMPLES AND HANDS-ON EXERCISES.

2. *MASTERING DATA SCIENCE WITH SQL SERVER 2017 AND R*

FOCUSED ON COMBINING SQL SERVER 2017 CAPABILITIES WITH R PROGRAMMING, THIS BOOK GUIDES READERS THROUGH ADVANCED DATA SCIENCE TECHNIQUES. IT EXPLORES HOW TO USE R SERVICES IN SQL SERVER FOR STATISTICAL ANALYSIS, DATA VISUALIZATION, AND PREDICTIVE MODELING. THE BOOK INCLUDES DETAILED CASE STUDIES TO DEMONSTRATE PRACTICAL APPLICATIONS IN BUSINESS INTELLIGENCE.

3. *PRACTICAL DATA SCIENCE USING SQL SERVER 2017 MACHINE LEARNING SERVICES*

THIS TITLE DIVES INTO THE MACHINE LEARNING SERVICES FEATURE OF SQL SERVER 2017, EXPLAINING HOW TO IMPLEMENT MACHINE LEARNING WORKFLOWS INSIDE THE DATABASE. IT COVERS INSTALLATION, SETUP, AND EXECUTION OF R AND PYTHON SCRIPTS WITHIN SQL SERVER. READERS LEARN TO CREATE SCALABLE DATA SCIENCE MODELS WITH SEAMLESS INTEGRATION TO EXISTING DATA INFRASTRUCTURE.

4. *DATA ANALYTICS WITH SQL SERVER 2017: A HANDS-ON GUIDE*

DESIGNED FOR DATA ANALYSTS AND SCIENTISTS, THIS BOOK TEACHES HOW TO PERFORM COMPREHENSIVE DATA ANALYTICS TASKS USING SQL SERVER 2017. IT INCLUDES DATA CLEANSING, AGGREGATION, AND ADVANCED QUERYING TECHNIQUES TAILORED FOR DATA SCIENCE PROJECTS. THE GUIDE ALSO EXPLORES INTEGRATION WITH VISUALIZATION TOOLS AND DEPLOYING ANALYTICS SOLUTIONS.

5. *SQL SERVER 2017 FOR DATA SCIENTISTS: BUILDING INTELLIGENT APPLICATIONS*

THIS BOOK FOCUSES ON USING SQL SERVER 2017 TO DEVELOP INTELLIGENT APPLICATIONS POWERED BY DATA SCIENCE MODELS. IT EXPLAINS HOW TO EMBED MACHINE LEARNING MODELS INTO SQL SERVER AND OPERATIONALIZE THEM FOR REAL-TIME ANALYTICS. READERS GAIN INSIGHTS INTO MODEL TRAINING, SCORING, AND LIFECYCLE MANAGEMENT WITHIN A SQL SERVER ENVIRONMENT.

6. *APPLIED DATA SCIENCE WITH SQL SERVER 2017 AND PYTHON*

COVERING THE INTEGRATION OF PYTHON IN SQL SERVER 2017, THIS BOOK OFFERS PRACTICAL EXAMPLES OF DATA SCIENCE WORKFLOWS INSIDE THE DATABASE. IT TEACHES HOW TO RUN PYTHON SCRIPTS FOR DATA MANIPULATION, MACHINE LEARNING, AND AUTOMATION USING SQL SERVER MACHINE LEARNING SERVICES. THE BOOK IS IDEAL FOR PYTHON DEVELOPERS TRANSITIONING TO DATABASE-CENTRIC DATA SCIENCE.

7. *SQL SERVER 2017 DATA SCIENCE COOKBOOK*

THIS COOKBOOK PROVIDES A COLLECTION OF READY-TO-USE RECIPES FOR COMMON DATA SCIENCE TASKS USING SQL SERVER

2017. FROM DATA PREPARATION TO ADVANCED MACHINE LEARNING ALGORITHMS, EACH RECIPE OFFERS STEP-BY-STEP INSTRUCTIONS AND SAMPLE CODE. IT'S A VALUABLE RESOURCE FOR PRACTITIONERS SEEKING QUICK SOLUTIONS AND BEST PRACTICES.

8. *BUILDING DATA SCIENCE SOLUTIONS WITH SQL SERVER 2017*

THIS COMPREHENSIVE GUIDE COVERS THE END-TO-END PROCESS OF CREATING DATA SCIENCE SOLUTIONS ON SQL SERVER 2017. IT INCLUDES DATA INGESTION, FEATURE ENGINEERING, MODEL DEVELOPMENT, AND DEPLOYMENT STRATEGIES. THE BOOK EMPHASIZES SCALABILITY AND PERFORMANCE OPTIMIZATION FOR ENTERPRISE-LEVEL DATA SCIENCE PROJECTS.

9. *DATA SCIENCE AND MACHINE LEARNING WITH SQL SERVER 2017*

THIS BOOK EXPLORES THE SYNERGY BETWEEN DATA SCIENCE METHODOLOGIES AND SQL SERVER 2017'S MACHINE LEARNING CAPABILITIES. IT EXPLAINS HOW TO IMPLEMENT CLASSIFICATION, REGRESSION, CLUSTERING, AND RECOMMENDATION SYSTEMS WITHIN THE DATABASE. READERS ARE GUIDED THROUGH PRACTICAL EXAMPLES AND BEST PRACTICES FOR EFFECTIVE MODEL INTEGRATION.

Hands On Data Science With Sql Server 2017

Find other PDF articles:

<https://a.comtex-nj.com/wwu16/pdf?dataid=JfM52-8105&title=selection-and-speciation-pogil-answer-key.pdf>

Hands-On Data Science with SQL Server 2017

Ebook Name: Unlocking Data Insights: A Practical Guide to Data Science with SQL Server 2017

Ebook Outline:

Introduction: What is Data Science? Why SQL Server 2017? Setting up your environment.

Chapter 1: Data Wrangling and Preprocessing with SQL Server: Data cleaning, transformation, and handling missing values using T-SQL.

Chapter 2: Exploratory Data Analysis (EDA) with SQL Server: Descriptive statistics, data visualization techniques within SQL Server, identifying patterns and anomalies.

Chapter 3: Machine Learning with SQL Server Machine Learning Services: Introduction to SQL Server ML Services, building and deploying simple predictive models (regression, classification).

Chapter 4: Advanced Analytics and Predictive Modeling: Working with more complex models, model evaluation, and feature engineering within the SQL Server environment.

Chapter 5: Data Visualization and Reporting: Creating insightful dashboards and reports using SQL Server Reporting Services (SSRS) or Power BI.

Chapter 6: Case Studies and Real-World Applications: Illustrative examples of data science projects using SQL Server 2017.

Conclusion: Future trends in data science and SQL Server, further learning resources.

Hands-On Data Science with SQL Server 2017: A Comprehensive Guide

The world is drowning in data. Turning this raw data into actionable insights is the core of data science, a field rapidly transforming industries across the board. While specialized languages like Python and R are often associated with data science, SQL Server 2017, and its subsequent versions, offers a powerful and often overlooked alternative, particularly for organizations already invested in the Microsoft ecosystem. This comprehensive guide will explore the capabilities of SQL Server 2017 for hands-on data science, demonstrating its effectiveness in tackling real-world problems.

1. Introduction: Setting the Stage for Data Science with SQL Server 2017

Data science, at its heart, is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data. This involves various stages: data collection, cleaning, exploration, modeling, and visualization. SQL Server 2017, with its robust database management system and integrated machine learning capabilities, provides a solid foundation for executing many of these steps within a single environment. This eliminates the need for data transfer between disparate systems, streamlining the entire data science workflow and reducing potential errors. This chapter will provide a high-level overview of data science methodologies and then focus on setting up the necessary environment for hands-on work with SQL Server 2017, including installation instructions and basic database management tasks.

2. Data Wrangling and Preprocessing with SQL Server: Taming the Data Beast

Raw data is rarely usable in its initial form. Data wrangling, or preprocessing, is a crucial step in any data science project. This involves cleaning, transforming, and preparing the data for analysis and modeling. SQL Server offers a powerful arsenal of T-SQL functions and techniques for performing these tasks effectively. This chapter focuses on:

Data Cleaning: Handling missing values (imputation or removal), dealing with outliers, and correcting inconsistencies in data. We'll explore various T-SQL functions like `ISNULL`, `CASE`, and `COALESCE` for handling missing data.

Data Transformation: Converting data types, creating new variables from existing ones, and standardizing data formats. We'll cover techniques like string manipulation, date and time functions, and aggregate functions.

Handling Missing Values: We'll delve into different strategies for handling missing values, weighing the pros and cons of imputation versus removal, and demonstrating how to implement these strategies using T-SQL.

By mastering these techniques, readers will be equipped to transform messy, raw data into a clean and consistent dataset ready for analysis.

3. Exploratory Data Analysis (EDA) with SQL Server: Uncovering Hidden Patterns

Exploratory Data Analysis (EDA) is an iterative process of exploring and visualizing data to gain insights, identify patterns, and formulate hypotheses. While tools like Python's Pandas and Matplotlib are frequently used for EDA, SQL Server provides surprisingly powerful capabilities for performing many EDA tasks directly within the database. This chapter will cover:

Descriptive Statistics: Calculating summary statistics (mean, median, standard deviation, etc.) using SQL Server's built-in functions.

Data Visualization: Generating basic visualizations directly within SQL Server using techniques like creating frequency distributions and histograms. We'll explore the limitations of SQL Server's built-in visualization capabilities and discuss how to integrate it with other visualization tools.

Pattern Identification and Anomaly Detection: Using SQL queries to identify trends, outliers, and anomalies in the data.

This chapter empowers readers to conduct efficient EDA, uncovering valuable information from their datasets before moving on to more complex modeling techniques.

4. Machine Learning with SQL Server Machine Learning Services: Building Predictive Models

SQL Server Machine Learning Services (formerly R Services) allow users to build and deploy machine learning models directly within the SQL Server environment. This integration eliminates the need for exporting data to external environments, streamlining the process and improving performance. This chapter will introduce:

Introduction to SQL Server ML Services: Setting up the environment and understanding the architecture of ML Services.

Building Simple Predictive Models: We'll build simple regression and classification models using popular algorithms like linear regression and logistic regression.

Model Deployment and Integration: Integrating the trained models into SQL Server for real-time prediction.

5. Advanced Analytics and Predictive Modeling: Scaling Up

Your Analysis

This chapter delves deeper into the capabilities of SQL Server ML Services, tackling more advanced techniques and challenges:

Working with More Complex Models: Exploring more sophisticated algorithms like support vector machines (SVMs), decision trees, and random forests.

Model Evaluation: Evaluating model performance using metrics such as accuracy, precision, recall, and F1-score. We'll also discuss cross-validation and other model selection techniques.

Feature Engineering: Creating new features from existing ones to improve model performance.

6. Data Visualization and Reporting: Communicating Your Findings

The insights gained through data analysis are only valuable if they can be effectively communicated. This chapter explores how to create compelling visualizations and reports using SQL Server's built-in tools and integrations:

SQL Server Reporting Services (SSRS): Creating interactive dashboards and reports using SSRS.

Power BI Integration: Leveraging Power BI's powerful visualization capabilities to create interactive dashboards and reports.

This chapter equips readers to present their findings clearly and concisely to stakeholders.

7. Case Studies and Real-World Applications: Putting it All Together

This chapter presents several case studies showcasing real-world applications of data science using SQL Server 2017. These examples will illustrate the practical implementation of the techniques covered throughout the ebook.

8. Conclusion: The Future of Data Science with SQL Server

This concluding chapter summarizes the key takeaways, discusses future trends in data science and SQL Server, and suggests resources for continued learning.

FAQs

1. What is the minimum SQL Server version required? While the ebook focuses on SQL Server 2017, many concepts apply to later versions as well. However, some features, especially within Machine Learning Services, may have evolved.
2. What programming languages are needed besides T-SQL? While T-SQL is the primary language, familiarity with Python or R can be beneficial for more advanced modeling and visualization.
3. Can I use SQL Server for all stages of a data science project? While SQL Server excels at data management, preprocessing, and certain modeling tasks, integrating it with other tools like Python or R may be necessary for more complex modeling or sophisticated visualizations.
4. How much experience with SQL is needed? A basic understanding of SQL is recommended.
5. Is this ebook suitable for beginners? Yes, the ebook is designed to be accessible to beginners, with clear explanations and practical examples.
6. What kind of datasets can I work with? SQL Server can handle various datasets, including relational databases, CSV files, and other structured data formats.
7. What are the limitations of using SQL Server for data science? SQL Server may not be the most efficient tool for extremely large datasets or highly complex machine learning models.
8. Are there any costs associated with using SQL Server? SQL Server has different licensing options, including free Express editions. Consider your needs and choose the appropriate licensing model.
9. Where can I find additional resources for learning? Microsoft's documentation, online tutorials, and community forums are excellent resources.

Related Articles

1. SQL Server 2017 Performance Tuning for Data Science: Optimizing SQL queries and database performance for faster data science workflows.
2. Advanced T-SQL Techniques for Data Wrangling: Mastering advanced T-SQL functions and techniques for efficient data cleaning and transformation.
3. Building Scalable Machine Learning Models with SQL Server: Techniques for building machine learning models that can handle large datasets and high volumes of requests.
4. Data Visualization Best Practices for Data Science Reports: Creating effective and visually appealing data visualizations to communicate insights effectively.

5. Comparing SQL Server Machine Learning Services with other ML Platforms: Evaluating the strengths and weaknesses of SQL Server ML Services compared to other platforms.
6. Integrating SQL Server with Python for Data Science: Techniques for combining the power of SQL Server with Python's rich data science libraries.
7. Case Study: Fraud Detection using SQL Server Machine Learning Services: A practical example of using SQL Server for building a fraud detection system.
8. Deploying SQL Server-Based Machine Learning Models to Production: Best practices for deploying and managing machine learning models in a production environment.
9. The Future of SQL Server in the Age of Big Data and Cloud Computing: Exploring how SQL Server is adapting to the evolving landscape of big data and cloud computing.

hands on data science with sql server 2017: Hands-On Data Science with SQL Server 2017
 Marek Chmel, Vladimír Mužný, 2018-11-29 Find, explore, and extract big data to transform into actionable insights Key FeaturesPerform end-to-end data analysis—from exploration to visualizationReal-world examples, tasks, and interview queries to be a proficient data scientistUnderstand how SQL is used for big data processing using HiveQL and SparkSQLBook Description SQL Server is a relational database management system that enables you to cover end-to-end data science processes using various inbuilt services and features. Hands-On Data Science with SQL Server 2017 starts with an overview of data science with SQL to understand the core tasks in data science. You will learn intermediate-to-advanced level concepts to perform analytical tasks on data using SQL Server. The book has a unique approach, covering best practices, tasks, and challenges to test your abilities at the end of each chapter. You will explore the ins and outs of performing various key tasks such as data collection, cleaning, manipulation, aggregations, and filtering techniques. As you make your way through the chapters, you will turn raw data into actionable insights by wrangling and extracting data from databases using T-SQL. You will get to grips with preparing and presenting data in a meaningful way, using Power BI to reveal hidden patterns. In the concluding chapters, you will work with SQL Server integration services to transform data into a useful format and delve into advanced examples covering machine learning concepts such as predictive analytics using real-world examples. By the end of this book, you will be in a position to handle the growing amounts of data and perform everyday activities that a data science professional performs. What you will learnUnderstand what data science is and how SQL Server is used for big data processingAnalyze incoming data with SQL queries and visualizationsCreate, train, and evaluate predictive modelsMake predictions using trained models and establish regular retraining coursesIncorporate data source querying into SQL ServerEnhance built-in T-SQL capabilities using SQLCLRVisualize data with Reporting Services, Power View, and Power BITransform data with R, Python, and AzureWho this book is for Hands-On Data Science with SQL Server 2017 is intended for data scientists, data analysts, and big data professionals who want to master their skills learning SQL and its applications. This book will be helpful even for beginners who want to build their career as data science professionals using the power of SQL Server 2017. Basic familiarity with SQL language will aid with understanding the concepts covered in this book.

hands on data science with sql server 2017: SQL Server 2017 Machine Learning Services with R
 Tomaz Kastrun, Julie Koesmarno, 2018-02-27 Develop and run efficient R scripts and predictive models for SQL Server 2017 Key Features Learn how you can combine the power of R and SQL Server 2017 to build efficient, cost-effective data science solutions Leverage the capabilities of R Services to perform advanced analytics—from data exploration to predictive modeling A quick primer with practical examples to help you get up- and- running with SQL Server 2017 Machine

Learning Services with R, as part of database solutions with continuous integration / continuous delivery. Book Description R Services was one of the most anticipated features in SQL Server 2016, improved significantly and rebranded as SQL Server 2017 Machine Learning Services. Prior to SQL Server 2016, many developers and data scientists were already using R to connect to SQL Server in siloed environments that left a lot to be desired, in order to do additional data analysis, superseding SSAS Data Mining or additional CLR programming functions. With R integrated within SQL Server 2017, these developers and data scientists can now benefit from its integrated, effective, efficient, and more streamlined analytics environment. This book gives you foundational knowledge and insights to help you understand SQL Server 2017 Machine Learning Services with R. First and foremost, the book provides practical examples on how to implement, use, and understand SQL Server and R integration in corporate environments, and also provides explanations and underlying motivations. It covers installing Machine Learning Services; maintaining, deploying, and managing code; and monitoring your services. Delving more deeply into predictive modeling and the RevoScaleR package, this book also provides insights into operationalizing code and exploring and visualizing data. To complete the journey, this book covers the new features in SQL Server 2017 and how they are compatible with R, amplifying their combined power. What you will learn Get an overview of SQL Server 2017 Machine Learning Services with R Manage SQL Server Machine Learning Services from installation to configuration and maintenance Handle and operationalize R code Explore RevoScaleR R algorithms and create predictive models Deploy, manage, and monitor database solutions with R Extend R with SQL Server 2017 features Explore the power of R for database administrators Who this book is for This book is for data analysts, data scientists, and database administrators with some or no experience in R but who are eager to easily deliver practical data science solutions in their day-to-day work (or future projects) using SQL Server.

hands on data science with sql server 2017: Learn T-SQL Querying Pedro Lopes, Pam Lahoud, 2019-05-03 Troubleshoot query performance issues, identify anti-patterns in code, and write efficient T-SQL queries Key Features Discover T-SQL functionalities and services that help you interact with relational databases Understand the roles, tasks, and responsibilities of a T-SQL developer Explore solutions for carrying out database querying tasks, database administration, and troubleshooting Book Description Transact-SQL (T-SQL) is Microsoft's proprietary extension to the SQL language used with Microsoft SQL Server and Azure SQL Database. This book will be a useful to learning the art of writing efficient T-SQL code in modern SQL Server versions as well as the Azure SQL Database. The book will get you started with query processing fundamentals to help you write powerful, performant T-SQL queries. You will then focus on query execution plans and leverage them for troubleshooting. In later chapters, you will explain how to identify various T-SQL patterns and anti-patterns. This will help you analyze execution plans to gain insights into current performance, and determine whether or not a query is scalable. You will also build diagnostic queries using dynamic management views (DMVs) and dynamic management functions (DMFs) to address various challenges in T-SQL execution. Next, you will work with the built-in tools of SQL Server to shorten the time taken to address query performance and scalability issues. In the concluding chapters, this will guide you through implementing various features, such as Extended Events, Query Store, and Query Tuning Assistant, using hands-on examples. By the end of the book, you will have developed the skills to determine query performance bottlenecks, avoid pitfalls, and discover the anti-patterns in use. What you will learn Use Query Store to understand and easily change query performance Recognize and eliminate bottlenecks that lead to slow performance Deploy quick fixes and long-term solutions to improve query performance Implement best practices to minimize performance risk using T-SQL Achieve optimal performance by ensuring careful query and index design Use the latest performance optimization features in SQL Server 2017 and SQL Server 2019 Protect query performance during upgrades to newer versions of SQL Server Who this book is for This book is for database administrators, database developers, data analysts, data scientists, and T-SQL practitioners who want to get started with writing T-SQL code and troubleshooting query performance issues with the help of practical examples. Previous knowledge

of T-SQL querying is not required to get started with this book.

hands on data science with sql server 2017: *Hands-On SQL Server 2019 Analysis Services*
Steven Hughes, 2020-10-22 Get up to speed with the new features added to Microsoft SQL Server 2019 Analysis Services and create models to support your business Key Features Explore tips and tricks to design, develop, and optimize end-to-end data analytics solutions using Microsoft's technologies Learn tabular modeling and multi-dimensional cube design development using real-world examples Implement Analysis Services to help you make productive business decisions Book Description SQL Server Analysis Services (SSAS) continues to be a leading enterprise-scale toolset, enabling customers to deliver data and analytics across large datasets with great performance. This book will help you understand MS SQL Server 2019's new features and improvements, especially when it comes to SSAS. First, you'll cover a quick overview of SQL Server 2019, learn how to choose the right analytical model to use, and understand their key differences. You'll then explore how to create a multi-dimensional model with SSAS and expand on that model with MDX. Next, you'll create and deploy a tabular model using Microsoft Visual Studio and Management Studio. You'll learn when and how to use both tabular and multi-dimensional model types, how to deploy and configure your servers to support them, and design principles that are relevant to each model. The book comes packed with tips and tricks to build measures, optimize your design, and interact with models using Excel and Power BI. All this will help you visualize data to gain useful insights and make better decisions. Finally, you'll discover practices and tools for securing and maintaining your models once they are deployed. By the end of this MS SQL Server book, you'll be able to choose the right model and build and deploy it to support the analytical needs of your business. What you will learn Determine the best analytical model using SSAS Cover the core aspects involved in MDX, including writing your first query Implement calculated tables and calculation groups (new in version 2019) in DAX Create and deploy tabular and multi-dimensional models on SQL 2019 Connect and create data visualizations using Excel and Power BI Implement row-level and other data security methods with tabular and multi-dimensional models Explore essential concepts and techniques to scale, manage, and optimize your SSAS solutions Who this book is for This Microsoft SQL Server book is for BI professionals and data analysts who are looking for a practical guide to creating and maintaining tabular and multi-dimensional models using SQL Server 2019 Analysis Services. A basic working knowledge of BI solutions such as Power BI and database querying is required.

hands on data science with sql server 2017: *Automated Machine Learning with Microsoft Azure*
Dennis Michael Sawyers, 2021-04-23 A practical, step-by-step guide to using Microsoft's AutoML technology on the Azure Machine Learning service for developers and data scientists working with the Python programming language Key Features Create, deploy, productionalize, and scale automated machine learning solutions on Microsoft Azure Improve the accuracy of your ML models through automatic data featurization and model training Increase productivity in your organization by using artificial intelligence to solve common problems Book Description Automated Machine Learning with Microsoft Azure will teach you how to build high-performing, accurate machine learning models in record time. It will equip you with the knowledge and skills to easily harness the power of artificial intelligence and increase the productivity and profitability of your business. Guided user interfaces (GUIs) enable both novices and seasoned data scientists to easily train and deploy machine learning solutions to production. Using a careful, step-by-step approach, this book will teach you how to use Azure AutoML with a GUI as well as the AzureML Python software development kit (SDK). First, you'll learn how to prepare data, train models, and register them to your Azure Machine Learning workspace. You'll then discover how to take those models and use them to create both automated batch solutions using machine learning pipelines and real-time scoring solutions using Azure Kubernetes Service (AKS). Finally, you will be able to use AutoML on your own data to not only train regression, classification, and forecasting models but also use them to solve a wide variety of business problems. By the end of this Azure book, you'll be able to show your business partners exactly how your ML models are making predictions through automatically

generated charts and graphs, earning their trust and respect. What you will learn

Understand how to train classification, regression, and forecasting ML algorithms with Azure AutoML

Prepare data for Azure AutoML to ensure smooth model training and deployment

Adjust AutoML configuration settings to make your models as accurate as possible

Determine when to use a batch-scoring solution versus a real-time scoring solution

Productionalize your AutoML and discover how to quickly deliver value

Create real-time scoring solutions with AutoML and Azure Kubernetes Service

Train a large number of AutoML models at once using the AzureML Python SDK

Who this book is for Data scientists, aspiring data scientists, machine learning engineers, or anyone interested in applying artificial intelligence or machine learning in their business will find this machine learning book useful. You need to have beginner-level knowledge of artificial intelligence and a technical background in computer science, statistics, or information technology before getting started. Familiarity with Python will help you implement the more advanced features found in the chapters, but even data analysts and SQL experts will be able to train ML models after finishing this book.

hands on data science with sql server 2017: *SQL Server 2017 Integration Services Cookbook*

Christian Cote, Matija Lah, Dejan Sarka, 2017-06-30

Harness the power of SQL Server 2017 Integration Services to build your data integration solutions with ease

About This Book

Acquaint yourself with all the newly introduced features in SQL Server 2017 Integration Services Program and extend your packages to enhance their functionality

This detailed, step-by-step guide covers everything you need to develop efficient data integration and data transformation solutions for your organization

Who This Book Is For

This book is ideal for software engineers, DW/ETL architects, and ETL developers who need to create a new, or enhance an existing, ETL implementation with SQL Server 2017 Integration Services. This book would also be good for individuals who develop ETL solutions that use SSIS and are keen to learn the new features and capabilities in SSIS 2017.

What You Will Learn

Understand the key components of an ETL solution using SQL Server 2016-2017 Integration Services

Design the architecture of a modern ETL solution

Have a good knowledge of the new capabilities and features added to Integration Services

Implement ETL solutions using Integration Services for both on-premises and Azure data

Improve the performance and scalability of an ETL solution

Enhance the ETL solution using a custom framework

Be able to work on the ETL solution with many other developers and have common design paradigms or techniques

Effectively use scripting to solve complex data issues

In Detail

SQL Server Integration Services is a tool that facilitates data extraction, consolidation, and loading options (ETL), SQL Server coding enhancements, data warehousing, and customizations. With the help of the recipes in this book, you'll gain complete hands-on experience of SSIS 2017 as well as the 2016 new features, design and development improvements including SCD, Tuning, and Customizations. At the start, you'll learn to install and set up SSIS as well other SQL Server resources to make optimal use of this Business Intelligence tools. We'll begin by taking you through the new features in SSIS 2016/2017 and implementing the necessary features to get a modern scalable ETL solution that fits the modern data warehouse. Through the course of chapters, you will learn how to design and build SSIS data warehouses packages using SQL Server Data Tools. Additionally, you'll learn to develop SSIS packages designed to maintain a data warehouse using the Data Flow and other control flow tasks. You'll also be demonstrated many recipes on cleansing data and how to get the end result after applying different transformations. Some real-world scenarios that you might face are also covered and how to handle various issues that you might face when designing your packages. At the end of this book, you'll get to know all the key concepts to perform data integration and transformation. You'll have explored on-premises Big Data integration processes to create a classic data warehouse, and will know how to extend the toolbox with custom tasks and transforms.

Style and approach

This cookbook follows a problem-solution approach and tackles all kinds of data integration scenarios by using the capabilities of SQL Server 2016 Integration Services. This book is well supplemented with screenshots, tips, and tricks. Each recipe focuses on a particular task and is written in a very easy-to-follow manner.

hands on data science with sql server 2017: Hands-On Data Science and Python Machine

Learning Frank Kane, 2017-07-31 This book covers the fundamentals of machine learning with Python in a concise and dynamic manner. It covers data mining and large-scale machine learning using Apache Spark. About This Book Take your first steps in the world of data science by understanding the tools and techniques of data analysis Train efficient Machine Learning models in Python using the supervised and unsupervised learning methods Learn how to use Apache Spark for processing Big Data efficiently Who This Book Is For If you are a budding data scientist or a data analyst who wants to analyze and gain actionable insights from data using Python, this book is for you. Programmers with some experience in Python who want to enter the lucrative world of Data Science will also find this book to be very useful, but you don't need to be an expert Python coder or mathematician to get the most from this book. What You Will Learn Learn how to clean your data and ready it for analysis Implement the popular clustering and regression methods in Python Train efficient machine learning models using decision trees and random forests Visualize the results of your analysis using Python's Matplotlib library Use Apache Spark's MLlib package to perform machine learning on large datasets In Detail Join Frank Kane, who worked on Amazon and IMDb's machine learning algorithms, as he guides you on your first steps into the world of data science. Hands-On Data Science and Python Machine Learning gives you the tools that you need to understand and explore the core topics in the field, and the confidence and practice to build and analyze your own machine learning models. With the help of interesting and easy-to-follow practical examples, Frank Kane explains potentially complex topics such as Bayesian methods and K-means clustering in a way that anybody can understand them. Based on Frank's successful data science course, Hands-On Data Science and Python Machine Learning empowers you to conduct data analysis and perform efficient machine learning using Python. Let Frank help you unearth the value in your data using the various data mining and data analysis techniques available in Python, and to develop efficient predictive models to predict future results. You will also learn how to perform large-scale machine learning on Big Data using Apache Spark. The book covers preparing your data for analysis, training machine learning models, and visualizing the final data analysis. Style and approach This comprehensive book is a perfect blend of theory and hands-on code examples in Python which can be used for your reference at any time.

hands on data science with sql server 2017: Hands-On Machine Learning with Azure Thomas K Abraham, Parashar Shah, Jen Stirrup, Lauri Lehman, Anindita Basak, 2018-10-31 Implement machine learning, cognitive services, and artificial intelligence solutions by leveraging Azure cloud technologies Key Features Learn advanced concepts in Azure ML and the Cortana Intelligence Suite architecture Explore ML Server using SQL Server and HDInsight capabilities Implement various tools in Azure to build and deploy machine learning models Book Description Implementing Machine learning (ML) and Artificial Intelligence (AI) in the cloud had not been possible earlier due to the lack of processing power and storage. However, Azure has created ML and AI services that are easy to implement in the cloud. Hands-On Machine Learning with Azure teaches you how to perform advanced ML projects in the cloud in a cost-effective way. The book begins by covering the benefits of ML and AI in the cloud. You will then explore Microsoft's Team Data Science Process to establish a repeatable process for successful AI development and implementation. You will also gain an understanding of AI technologies available in Azure and the Cognitive Services APIs to integrate them into bot applications. This book lets you explore prebuilt templates with Azure Machine Learning Studio and build a model using canned algorithms that can be deployed as web services. The book then takes you through a preconfigured series of virtual machines in Azure targeted at AI development scenarios. You will get to grips with the ML Server and its capabilities in SQL and HDInsight. In the concluding chapters, you'll integrate patterns with other non-AI services in Azure. By the end of this book, you will be fully equipped to implement smart cognitive actions in your models. What you will learn Discover the benefits of leveraging the cloud for ML and AI Use Cognitive Services APIs to build intelligent bots Build a model using canned algorithms from Microsoft and deploy it as a web service Deploy virtual machines in AI development scenarios Apply R, Python, SQL Server, and Spark in Azure Build and deploy deep learning solutions

with CNTK, MMLSpark, and TensorFlowImplement model retraining in IoT, Streaming, and Blockchain solutionsExplore best practices for integrating ML and AI functions with ADLA and logic appsWho this book is for If you are a data scientist or developer familiar with Azure ML and cognitive services and want to create smart models and make sense of data in the cloud, this book is for you. You'll also find this book useful if you want to bring powerful machine learning services into your cloud applications. Some experience with data manipulation and processing, using languages like SQL, Python, and R, will aid in understanding the concepts covered in this book

hands on data science with sql server 2017: Data Analysis Using SQL and Excel Gordon S. Linoff, 2010-09-16 Useful business analysis requires you to effectively transform data into actionable information. This book helps you use SQL and Excel to extract business information from relational databases and use that data to define business dimensions, store transactions about customers, produce results, and more. Each chapter explains when and why to perform a particular type of business analysis in order to obtain useful results, how to design and perform the analysis using SQL and Excel, and what the results should look like.

hands on data science with sql server 2017: Building Data Science Teams DJ Patil, 2011-09-15 As data science evolves to become a business necessity, the importance of assembling a strong and innovative data teams grows. In this in-depth report, data scientist DJ Patil explains the skills, perspectives, tools and processes that position data science teams for success. Topics include: What it means to be data driven. The unique roles of data scientists. The four essential qualities of data scientists. Patil's first-hand experience building the LinkedIn data science team.

hands on data science with sql server 2017: SQL Queries for Mere Mortals John L. Viescas, Michael James Hernandez, 2014 The #1 Easy, Common-Sense Guide to SQL Queries--Updated for Today's Databases, Standards, and Challenges SQL Queries for Mere Mortals® has earned worldwide praise as the clearest, simplest tutorial on writing effective SQL queries. The authors have updated this hands-on classic to reflect new SQL standards and database applications and teach valuable new techniques. Step by step, John L. Viescas and Michael J. Hernandez guide you through creating reliable queries for virtually any modern SQL-based database. They demystify all aspects of SQL query writing, from simple data selection and filtering to joining multiple tables and modifying sets of data. Three brand-new chapters teach you how to solve a wide range of challenging SQL problems. You'll learn how to write queries that apply multiple complex conditions on one table, perform sophisticated logical evaluations, and think outside the box using unlinked tables. Coverage includes -- Getting started: understanding what relational databases are, and ensuring that your database structures are sound -- SQL basics: using SELECT statements, creating expressions, sorting information with ORDER BY, and filtering data using WHERE -- Summarizing and grouping data with GROUP BY and HAVING clauses -- Drawing data from multiple tables: using INNER JOIN, OUTER JOIN, and UNION operators, and working with subqueries -- Modifying data sets with UPDATE, INSERT, and DELETE statements Advanced queries: complex NOT and AND, conditions, if-then-else using CASE, unlinked tables, driver tables, and more Practice all you want with downloadable sample databases for today's versions of Microsoft Office Access, Microsoft SQL Server, and the open source MySQL database. Whether you're a DBA, developer, user, or student, there's no better way to master SQL. informit.com/awforMereMortals.com

hands on data science with sql server 2017: Proceedings of 4th International Conference on BigData Analysis and Data Mining 2017 ConferenceSeries, September 07-08, 2017 Paris, France Key Topics : Cloud computing, Forecasting from Big Data, Optimization and Big Data, New visualization techniques, Social network analysis, Search and data mining, Complexity and Algorithms, Open Data, ETL (Extract, Transform and Load), OLAP Technologies, Big Data Algorithm, Data Mining Analysis, Kernel Methods, Frequent Pattern Mining, Clustering, Data Privacy and Ethics, Big Data Technologies, Business Analytics, Data Mining Methods and Algorithms, Data Mining Tasks and Processes, Data Mining Applications in Science, Engineering, Healthcare and Medicine, Big Data Applications, Data Mining Tools and Software, Data

Warehousing, Artificial Intelligence,

hands on data science with sql server 2017: SQL Server 2017 Developer's Guide William Durkin, Miloš Radivojević, Dejan Sarka, 2018-03-16 Build smarter and efficient database application systems for your organization with SQL Server 2017 Key Features Build database applications by using the development features of SQL Server 2017 Work with temporal tables to get information stored in a table at any time Use adaptive querying to enhance the performance of your queries Book Description Microsoft SQL Server 2017 is the next big step in the data platform history of Microsoft as it brings in the power of R and Python for machine learning and containerization-based deployment on Windows and Linux. Compared to its predecessor, SQL Server 2017 has evolved into Machine Learning with R services for statistical analysis and Python packages for analytical processing. This book prepares you for more advanced topics by starting with a quick introduction to SQL Server 2017's new features and a recapitulation of the possibilities you may have already explored with previous versions of SQL Server. The next part introduces you to enhancements in the Transact-SQL language and new database engine capabilities and then switches to a completely new technology inside SQL Server: JSON support. We also take a look at the Stretch database, security enhancements, and temporal tables. Furthermore, the book focuses on implementing advanced topics, including Query Store, columnstore indexes, and In-Memory OLTP. Towards the end of the book, you'll be introduced to R and how to use the R language with Transact-SQL for data exploration and analysis. You'll also learn to integrate Python code in SQL Server and graph database implementations along with deployment options on Linux and SQL Server in containers for development and testing. By the end of this book, you will have the required information to design efficient, high-performance database applications without any hassle. What you will learn Explore the new development features introduced in SQL Server 2017 Identify opportunities for In-Memory OLTP technology Use columnstore indexes to get storage and performance improvements Exchange JSON data between applications and SQL Server Use the new security features to encrypt or mask the data Control the access to the data on the row levels Discover the potential of R and Python integration Model complex relationships with the graph databases in SQL Server 2017 Who this book is for Database developers and solution architects looking to design efficient database applications using SQL Server 2017 will find this book very useful. In addition, this book will be valuable to advanced analysis practitioners and business intelligence developers. Database consultants dealing with performance tuning will get a lot of useful information from this book as well. Some basic understanding of database concepts and T-SQL is required to get the best out of this book.

hands on data science with sql server 2017: Hands-On Cloud Solutions with Azure Greg Leonardo, 2018-10-31 Design effective Azure architecture and transform your IT business solutions Key Features Develop a resilient and robust cloud environment Deploy and manage cost-effective and highly available solutions on your public cloud Design and implement enterprise-level cloud solutions Book Description Azure provides cloud-based solutions to support your business demands. Building and running solutions on Azure will help your business maximize the return on investment and minimize the total cost of ownership. Hands-On Cloud Solutions with Azure focuses on addressing the architectural decisions that usually arise when you design or migrate a solution to Microsoft Azure. You will start by designing the building blocks of infrastructure solution on Azure, such as Azure compute, storage, and networking, followed by exploring the database options it offers. You will get to grips with designing scalable web and mobile solutions and understand where to host your Active Directory and Identity Solution. Moving on, you'll learn how to extend DevOps to Azure. You will also benefit from some exciting services that enable extremely smooth operations and streamlined DevOps between on-premises and cloud. The book will help you to design a secure environment for your solution, on both the Cloud and hybrid. Toward the end, you'll see how to manage and monitor cloud and hybrid solutions. By the end of this book, you will be armed with all the tools and knowledge you need to properly plan and design your solutions on Azure, whether it's for a brand new project or migration project. What you will learn Get started with Azure by

understanding tenants, subs, and resource groups Decide whether to “lift and shift” or migrate apps Plan and architect solutions in Azure Build ARM templates for Azure resources Develop and deploy solutions in Azure Understand how to monitor and support your application with Azure Make your life easier with Azure best practices and tips Who this book is for If you’re an IT consultant, developer, or solutions architect looking to design effective solutions for your organization, this book is for you. Some knowledge of cloud computing will assist with understanding the key concepts covered in this book.

hands on data science with sql server 2017: SQL for Data Analysis Cathy Tanimura, 2021-09-09 With the explosion of data, computing power, and cloud data warehouses, SQL has become an even more indispensable tool for the savvy analyst or data scientist. This practical book reveals new and hidden ways to improve your SQL skills, solve problems, and make the most of SQL as part of your workflow. You'll learn how to use both common and exotic SQL functions such as joins, window functions, subqueries, and regular expressions in new, innovative ways--as well as how to combine SQL techniques to accomplish your goals faster, with understandable code. If you work with SQL databases, this is a must-have reference. Learn the key steps for preparing your data for analysis Perform time series analysis using SQL's date and time manipulations Use cohort analysis to investigate how groups change over time Use SQL's powerful functions and operators for text analysis Detect outliers in your data and replace them with alternate values Establish causality using experiment analysis, also known as A/B testing

hands on data science with sql server 2017: Scala and Spark for Big Data Analytics Md. Rezaul Karim, Sridhar Alla, 2017-07-25 Harness the power of Scala to program Spark and analyze tonnes of data in the blink of an eye! About This Book Learn Scala's sophisticated type system that combines Functional Programming and object-oriented concepts Work on a wide array of applications, from simple batch jobs to stream processing and machine learning Explore the most common as well as some complex use-cases to perform large-scale data analysis with Spark Who This Book Is For Anyone who wishes to learn how to perform data analysis by harnessing the power of Spark will find this book extremely useful. No knowledge of Spark or Scala is assumed, although prior programming experience (especially with other JVM languages) will be useful to pick up concepts quicker. What You Will Learn Understand object-oriented & functional programming concepts of Scala In-depth understanding of Scala collection APIs Work with RDD and DataFrame to learn Spark's core abstractions Analysing structured and unstructured data using SparkSQL and GraphX Scalable and fault-tolerant streaming application development using Spark structured streaming Learn machine-learning best practices for classification, regression, dimensionality reduction, and recommendation system to build predictive models with widely used algorithms in Spark MLlib & ML Build clustering models to cluster a vast amount of data Understand tuning, debugging, and monitoring Spark applications Deploy Spark applications on real clusters in Standalone, Mesos, and YARN In Detail Scala has been observing wide adoption over the past few years, especially in the field of data science and analytics. Spark, built on Scala, has gained a lot of recognition and is being used widely in productions. Thus, if you want to leverage the power of Scala and Spark to make sense of big data, this book is for you. The first part introduces you to Scala, helping you understand the object-oriented and functional programming concepts needed for Spark application development. It then moves on to Spark to cover the basic abstractions using RDD and DataFrame. This will help you develop scalable and fault-tolerant streaming applications by analyzing structured and unstructured data using SparkSQL, GraphX, and Spark structured streaming. Finally, the book moves on to some advanced topics, such as monitoring, configuration, debugging, testing, and deployment. You will also learn how to develop Spark applications using SparkR and PySpark APIs, interactive data analytics using Zeppelin, and in-memory data processing with Alluxio. By the end of this book, you will have a thorough understanding of Spark, and you will be able to perform full-stack data analytics with a feel that no amount of data is too big. Style and approach Filled with practical examples and use cases, this book will not only help you get up and running with Spark, but will also take you farther down the road to becoming a data scientist.

hands on data science with sql server 2017: Python Data Science Handbook Jake VanderPlas, 2016-11-21 For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the Python Data Science Handbook do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you'll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine learning algorithms

hands on data science with sql server 2017: End-to-End Data Science with SAS James Gearheart, 2020-06-26 Learn data science concepts with real-world examples in SAS! End-to-End Data Science with SAS: A Hands-On Programming Guide provides clear and practical explanations of the data science environment, machine learning techniques, and the SAS programming knowledge necessary to develop machine learning models in any industry. The book covers concepts including understanding the business need, creating a modeling data set, linear regression, parametric classification models, and non-parametric classification models. Real-world business examples and example code are used to demonstrate each process step-by-step. Although a significant amount of background information and supporting mathematics are presented, the book is not structured as a textbook, but rather it is a user's guide for the application of data science and machine learning in a business environment. Readers will learn how to think like a data scientist, wrangle messy data, choose a model, and evaluate the model's effectiveness. New data scientists or professionals who want more experience with SAS will find this book to be an invaluable reference. Take your data science career to the next level by mastering SAS programming for machine learning models.

hands on data science with sql server 2017: A Practical Hands-on Approach to Database Forensics Nhien-An Le-Khac, Kim-Kwang Raymond Choo, 2022-10-21 Adopting an experimental learning approach, this book describes a practical forensic process to acquire and analyze databases from a given device and/or application. Databases hold important, sensitive, and/or confidential information and are a crucial source of evidence in any digital investigation. This also reinforces the importance of keeping up to date on the cyber-threat landscape as well as any associated database forensic challenges and approaches. The book also guides cyber-forensic researchers, educators, and practitioners through the process of conducting database forensics and investigations on mobile devices, Internet of Things (IoT) devices, web browsers, and end-to-end encrypted instant messaging applications. Given the fast-changing database forensics landscape, this book will be of interest to researchers, educators, and practitioners in the field, as well as students who want to learn about the database investigation.

hands on data science with sql server 2017: Data Science for Public Policy Jeffrey C. Chen, Edward A. Rubin, Gary J. Cornwall, 2021-09-01 This textbook presents the essential tools and core concepts of data science to public officials, policy analysts, and economists among others in order to further their application in the public sector. An expansion of the quantitative economics frameworks presented in policy and business schools, this book emphasizes the process of asking relevant questions to inform public policy. Its techniques and approaches emphasize data-driven practices, beginning with the basic programming paradigms that occupy the majority of an analyst's time and advancing to the practical applications of statistical learning and machine learning. The text considers two divergent, competing perspectives to support its applications, incorporating techniques from both causal inference and prediction. Additionally, the book includes open-sourced

data as well as live code, written in R and presented in notebook form, which readers can use and modify to practice working with data.

hands on data science with sql server 2017: Modern Data Science with R Benjamin S. Baumer, Daniel T. Kaplan, Nicholas J. Horton, 2021-03-31 From a review of the first edition: Modern Data Science with R... is rich with examples and is guided by a strong narrative voice. What's more, it presents an organizing framework that makes a convincing argument that data science is a course distinct from applied statistics (The American Statistician). Modern Data Science with R is a comprehensive data science textbook for undergraduates that incorporates statistical and computational thinking to solve real-world data problems. Rather than focus exclusively on case studies or programming syntax, this book illustrates how statistical programming in the state-of-the-art R/RStudio computing environment can be leveraged to extract meaningful information from a variety of data in the service of addressing compelling questions. The second edition is updated to reflect the growing influence of the tidyverse set of packages. All code in the book has been revised and styled to be more readable and easier to understand. New functionality from packages like sf, purrr, tidymodels, and tidytext is now integrated into the text. All chapters have been revised, and several have been split, re-organized, or re-imagined to meet the shifting landscape of best practice.

hands on data science with sql server 2017: T-SQL Querying Itzik Ben-Gan, Adam Machanic, Dejan Sarka, Kevin Farlee, 2015-02-17 T-SQL insiders help you tackle your toughest queries and query-tuning problems Squeeze maximum performance and efficiency from every T-SQL query you write or tune. Four leading experts take an in-depth look at T-SQL's internal architecture and offer advanced practical techniques for optimizing response time and resource usage. Emphasizing a correct understanding of the language and its foundations, the authors present unique solutions they have spent years developing and refining. All code and techniques are fully updated to reflect new T-SQL enhancements in Microsoft SQL Server 2014 and SQL Server 2012. Write faster, more efficient T-SQL code: Move from procedural programming to the language of sets and logic Master an efficient top-down tuning methodology Assess algorithmic complexity to predict performance Compare data aggregation techniques, including new grouping sets Efficiently perform data-analysis calculations Make the most of T-SQL's optimized bulk import tools Avoid date/time pitfalls that lead to buggy, poorly performing code Create optimized BI statistical queries without additional software Use programmable objects to accelerate queries Unlock major performance improvements with In-Memory OLTP Master useful and elegant approaches to manipulating graphs About This Book For experienced T-SQL practitioners Includes coverage updated from Inside Microsoft SQL Server 2008 T-SQL Querying and Inside Microsoft SQL Server 2008 T-SQL Programming Valuable to developers, DBAs, BI professionals, and data scientists Covers many MCSE 70-464 and MCSA/MCSE 70-461 exam topics

hands on data science with sql server 2017: Encyclopedia of Data Science and Machine Learning Wang, John, 2023-01-20 Big data and machine learning are driving the Fourth Industrial Revolution. With the age of big data upon us, we risk drowning in a flood of digital data. Big data has now become a critical part of both the business world and daily life, as the synthesis and synergy of machine learning and big data has enormous potential. Big data and machine learning are projected to not only maximize citizen wealth, but also promote societal health. As big data continues to evolve and the demand for professionals in the field increases, access to the most current information about the concepts, issues, trends, and technologies in this interdisciplinary area is needed. The Encyclopedia of Data Science and Machine Learning examines current, state-of-the-art research in the areas of data science, machine learning, data mining, and more. It provides an international forum for experts within these fields to advance the knowledge and practice in all facets of big data and machine learning, emphasizing emerging theories, principals, models, processes, and applications to inspire and circulate innovative findings into research, business, and communities. Covering topics such as benefit management, recommendation system analysis, and global software development, this expansive reference provides a dynamic resource for data scientists, data

analysts, computer scientists, technical managers, corporate executives, students and educators of higher education, government officials, researchers, and academicians.

hands on data science with sql server 2017: *Apache Spark 2.x Machine Learning Cookbook* Siamak Amirghodsi, Meenakshi Rajendran, Broderick Hall, Shuen Mei, 2017-09-22 Simplify machine learning model implementations with Spark About This Book Solve the day-to-day problems of data science with Spark This unique cookbook consists of exciting and intuitive numerical recipes Optimize your work by acquiring, cleaning, analyzing, predicting, and visualizing your data Who This Book Is For This book is for Scala developers with a fairly good exposure to and understanding of machine learning techniques, but lack practical implementations with Spark. A solid knowledge of machine learning algorithms is assumed, as well as hands-on experience of implementing ML algorithms with Scala. However, you do not need to be acquainted with the Spark ML libraries and ecosystem. What You Will Learn Get to know how Scala and Spark go hand-in-hand for developers when developing ML systems with Spark Build a recommendation engine that scales with Spark Find out how to build unsupervised clustering systems to classify data in Spark Build machine learning systems with the Decision Tree and Ensemble models in Spark Deal with the curse of high-dimensionality in big data using Spark Implement Text analytics for Search Engines in Spark Streaming Machine Learning System implementation using Spark In Detail Machine learning aims to extract knowledge from data, relying on fundamental concepts in computer science, statistics, probability, and optimization. Learning about algorithms enables a wide range of applications, from everyday tasks such as product recommendations and spam filtering to cutting edge applications such as self-driving cars and personalized medicine. You will gain hands-on experience of applying these principles using Apache Spark, a resilient cluster computing system well suited for large-scale machine learning tasks. This book begins with a quick overview of setting up the necessary IDEs to facilitate the execution of code examples that will be covered in various chapters. It also highlights some key issues developers face while working with machine learning algorithms on the Spark platform. We progress by uncovering the various Spark APIs and the implementation of ML algorithms with developing classification systems, recommendation engines, text analytics, clustering, and learning systems. Toward the final chapters, we'll focus on building high-end applications and explain various unsupervised methodologies and challenges to tackle when implementing with big data ML systems. Style and approach This book is packed with intuitive recipes supported with line-by-line explanations to help you understand how to optimize your work flow and resolve problems when working with complex data modeling tasks and predictive algorithms. This is a valuable resource for data scientists and those working on large scale data projects.

hands on data science with sql server 2017: *Tabular Modeling with SQL Server 2016 Analysis Services Cookbook* Derek Wilson, 2017-01-30 Expert tabular modeling techniques for building and deploying cutting-edge business analytical reporting solutions About This Book Build and deploy Tabular Model projects from relational data sources Leverage DAX and create high-performing calculated fields and measures Create ad-hoc reports based on a Tabular Model solution Useful tips to monitor and optimize your tabular solutions Who This Book Is For This book is for SQL BI professionals and Architects who want to exploit the full power of the new Tabular models in Analysis Services. Some knowledge of previous versions of Analysis services would be helpful but is not essential. What You Will Learn Learn all about Tabular services mode and how it speeds up development Build solutions using sample datasets Explore built-in actions and transitions in SSAS 2016 Implement row-column, and role-based security in a Tabular Data model Realize the benefits of in-memory and DirectQuery deployment modes Get up to date with the new features added to SQL Server 2016 Analysis Services Optimize Data Models and Relationships Usage In Detail SQL Server Analysis Service (SSAS) has been widely used across multiple businesses to build smart online analytical reporting solutions. It includes two different types of modeling for analysis services: Tabular and Multi Dimensional. This book covers Tabular modeling, which uses tables and relationships with a fast in-memory engine to provide state of the art compression algorithms and

query performance. The book begins by quickly taking you through the concepts required to model tabular data and set up the necessary tools and services. As you learn to create tabular models using tools such as Excel and Power View, you'll be shown various strategies to deploy your model on the server and choose a query mode (In-memory or DirectQuery) that best suits your reporting needs. You'll also learn how to implement key and newly introduced DAX functions to create calculated columns and measures for your model data. Last but not least, you'll be shown techniques that will help you administer and secure your BI implementation along with some widely used tips and tricks to optimize your reporting solution. By the end of this book, you'll have gained hands-on experience with the powerful new features that have been added to Tabular models in SSAS 2016 and you'll be able to improve user satisfaction with faster reports and analytical queries. Style and approach This book takes a practical, recipe-based approach where each recipe lists the steps to address or implement a solution. You will be provided with several approaches to creating a business intelligence semantic model using analysis services.

hands on data science with sql server 2017: Mastering Spark for Data Science Andrew Morgan, Antoine Amend, David George, Matthew Hallett, 2017-03-29 Master the techniques and sophisticated analytics used to construct Spark-based solutions that scale to deliver production-grade data science products About This Book Develop and apply advanced analytical techniques with Spark Learn how to tell a compelling story with data science using Spark's ecosystem Explore data at scale and work with cutting edge data science methods Who This Book Is For This book is for those who have beginner-level familiarity with the Spark architecture and data science applications, especially those who are looking for a challenge and want to learn cutting edge techniques. This book assumes working knowledge of data science, common machine learning methods, and popular data science tools, and assumes you have previously run proof of concept studies and built prototypes. What You Will Learn Learn the design patterns that integrate Spark into industrialized data science pipelines See how commercial data scientists design scalable code and reusable code for data science services Explore cutting edge data science methods so that you can study trends and causality Discover advanced programming techniques using RDD and the DataFrame and Dataset APIs Find out how Spark can be used as a universal ingestion engine tool and as a web scraper Practice the implementation of advanced topics in graph processing, such as community detection and contact chaining Get to know the best practices when performing Extended Exploratory Data Analysis, commonly used in commercial data science teams Study advanced Spark concepts, solution design patterns, and integration architectures Demonstrate powerful data science pipelines In Detail Data science seeks to transform the world using data, and this is typically achieved through disrupting and changing real processes in real industries. In order to operate at this level you need to build data science solutions of substance –solutions that solve real problems. Spark has emerged as the big data platform of choice for data scientists due to its speed, scalability, and easy-to-use APIs. This book deep dives into using Spark to deliver production-grade data science solutions. This process is demonstrated by exploring the construction of a sophisticated global news analysis service that uses Spark to generate continuous geopolitical and current affairs insights. You will learn all about the core Spark APIs and take a comprehensive tour of advanced libraries, including Spark SQL, Spark Streaming, MLlib, and more. You will be introduced to advanced techniques and methods that will help you to construct commercial-grade data products. Focusing on a sequence of tutorials that deliver a working news intelligence service, you will learn about advanced Spark architectures, how to work with geographic data in Spark, and how to tune Spark algorithms so they scale linearly. Style and approach This is an advanced guide for those with beginner-level familiarity with the Spark architecture and working with Data Science applications. Mastering Spark for Data Science is a practical tutorial that uses core Spark APIs and takes a deep dive into advanced libraries including: Spark SQL, visual streaming, and MLlib. This book expands on titles like: Machine Learning with Spark and Learning Spark. It is the next learning curve for those comfortable with Spark and looking to improve their skills.

hands on data science with sql server 2017: SQL Cookbook Anthony Molinaro, 2006 A

guide to SQL covers such topics as retrieving records, metadata queries, working with strings, data arithmetic, date manipulation, reporting and warehousing, and hierarchical queries.

hands on data science with sql server 2017: Hands-On Data Science with R Vitor Bianchi Lanzetta, Nataraj Dasgupta, Ricardo Anjoletto Farias, 2018-11-30 A hands-on guide for professionals to perform various data science tasks in R Key Features Explore the popular R packages for data science Use R for efficient data mining, text analytics and feature engineering Become a thorough data science professional with the help of hands-on examples and use-cases in R Book Description R is the most widely used programming language, and when used in association with data science, this powerful combination will solve the complexities involved with unstructured datasets in the real world. This book covers the entire data science ecosystem for aspiring data scientists, right from zero to a level where you are confident enough to get hands-on with real-world data science problems. The book starts with an introduction to data science and introduces readers to popular R libraries for executing data science routine tasks. This book covers all the important processes in data science such as data gathering, cleaning data, and then uncovering patterns from it. You will explore algorithms such as machine learning algorithms, predictive analytical models, and finally deep learning algorithms. You will learn to run the most powerful visualization packages available in R so as to ensure that you can easily derive insights from your data. Towards the end, you will also learn how to integrate R with Spark and Hadoop and perform large-scale data analytics without much complexity. What you will learn Understand the R programming language and its ecosystem of packages for data science Obtain and clean your data before processing Master essential exploratory techniques for summarizing data Examine various machine learning prediction, models Explore the H2O analytics platform in R for deep learning Apply data mining techniques to available datasets Work with interactive visualization packages in R Integrate R with Spark and Hadoop for large-scale data analytics Who this book is for If you are a budding data scientist keen to learn about the popular pandas library, or a Python developer looking to step into the world of data analysis, this book is the ideal resource you need to get started. Some programming experience in Python will be helpful to get the most out of this course

hands on data science with sql server 2017: The Data Science Design Manual Steven S. Skiena, 2017-07-01 This engaging and clearly written textbook/reference provides a must-have introduction to the rapidly emerging interdisciplinary field of data science. It focuses on the principles fundamental to becoming a good data scientist and the key skills needed to build systems for collecting, analyzing, and interpreting data. The Data Science Design Manual is a source of practical insights that highlights what really matters in analyzing data, and provides an intuitive understanding of how these core concepts can be used. The book does not emphasize any particular programming language or suite of data-analysis tools, focusing instead on high-level discussion of important design principles. This easy-to-read text ideally serves the needs of undergraduate and early graduate students embarking on an "Introduction to Data Science" course. It reveals how this discipline sits at the intersection of statistics, computer science, and machine learning, with a distinct heft and character of its own. Practitioners in these and related fields will find this book perfect for self-study as well. Additional learning tools: Contains "War Stories," offering perspectives on how data science applies in the real world Includes "Homework Problems," providing a wide range of exercises and projects for self-study Provides a complete set of lecture slides and online video lectures at www.data-manual.com Provides "Take-Home Lessons," emphasizing the big-picture concepts to learn from each chapter Recommends exciting "Kaggle Challenges" from the online platform Kaggle Highlights "False Starts," revealing the subtle reasons why certain approaches fail Offers examples taken from the data science television show "The Quant Shop" (www.quant-shop.com)

hands on data science with sql server 2017: SQL Server 2017 Administrator's Guide Marek Chmel, Vladimir Muzny, 2017-12-12 Implement and administer successful database solution with SQL Server 2017 About This Book Master the required skills to successfully set up, administer, and maintain your SQL Server 2017 database solution Design and configure, manage, and secure a

rock-solid SQL server Comprehensive guide in keeping your SQL server disaster proof and all-time availability Who This Book Is For This book targets database administrators with an interest in SQL Server 2017 administration. Readers are expected to have some experience with previous SQL Server versions. What You Will Learn Learn about the new features of SQL Server 2017 and how to implement them Build a stable and fast SQL Server environment Fix performance issues by optimizing queries and making use of indexes Perform a health check of an existing troublesome database environment Design and use an optimal database management strategy Implement efficient backup and recovery techniques in-line with security policies Combine SQL Server 2017 and Azure and manage your solution by various automation techniques Perform data migration, cluster upgradation and server consolidation In Detail Take advantage of the real power of SQL Server 2017 with all its new features, in addition to covering core database administration tasks. This book will give you a competitive advantage by helping you quickly learn how to design, manage, and secure your database solution. You will learn how to set up your SQL Server and configure new (and existing) environments for optimal use. After covering the designing aspect, the book delves into performance-tuning aspects by teaching you how to effectively use indexes. The book will also teach you about certain choices that need to be made about backups and how to implement a rock-solid security policy and keep your environment healthy. Finally, you will learn about the techniques you should use when things go wrong, and other important topics - such as migration, upgrading, and consolidation - are covered in detail. Integration with Azure is also covered in depth. Whether you are an administrator or thinking about entering the field, this book will provide you with all the skills you need to successfully create, design, and deploy databases using SQL Server 2017. Style and approach A comprehensive guide for database professionals, covering a wide range of topics from installation, maintenance, and configuration to managing systems for operational efficiency and high availability; best practices for maintaining a highly reliable database solution are also supplied from industry experts.

hands on data science with sql server 2017: Modern Data Science with R Benjamin S. Baumer, Daniel T. Kaplan, Nicholas J. Horton, 2017-03-16 Modern Data Science with R is a comprehensive data science textbook for undergraduates that incorporates statistical and computational thinking to solve real-world problems with data. Rather than focus exclusively on case studies or programming syntax, this book illustrates how statistical programming in the state-of-the-art R/RStudio computing environment can be leveraged to extract meaningful information from a variety of data in the service of addressing compelling statistical questions. Contemporary data science requires a tight integration of knowledge from statistics, computer science, mathematics, and a domain of application. This book will help readers with some background in statistics and modest prior experience with coding develop and practice the appropriate skills to tackle complex data science projects. The book features a number of exercises and has a flexible organization conducive to teaching a variety of semester courses.

hands on data science with sql server 2017: Data Science For Dummies Lillian Pierson, 2021-08-20 Monetize your company's data and data science expertise without spending a fortune on hiring independent strategy consultants to help What if there was one simple, clear process for ensuring that all your company's data science projects achieve a high a return on investment? What if you could validate your ideas for future data science projects, and select the one idea that's most prime for achieving profitability while also moving your company closer to its business vision? There is. Industry-acclaimed data science consultant, Lillian Pierson, shares her proprietary STAR Framework - A simple, proven process for leading profit-forming data science projects. Not sure what data science is yet? Don't worry! Parts 1 and 2 of Data Science For Dummies will get all the bases covered for you. And if you're already a data science expert? Then you really won't want to miss the data science strategy and data monetization gems that are shared in Part 3 onward throughout this book. Data Science For Dummies demonstrates: The only process you'll ever need to lead profitable data science projects Secret, reverse-engineered data monetization tactics that no one's talking about The shocking truth about how simple natural language processing can be How to

beat the crowd of data professionals by cultivating your own unique blend of data science expertise. Whether you're new to the data science field or already a decade in, you're sure to learn something new and incredibly valuable from *Data Science For Dummies*. Discover how to generate massive business wins from your company's data by picking up your copy today.

hands on data science with sql server 2017: Introduction to Data Science Laura Igual, Santi Seguí, 2017-02-22 This accessible and classroom-tested textbook/reference presents an introduction to the fundamentals of the emerging and interdisciplinary field of data science. The coverage spans key concepts adopted from statistics and machine learning, useful techniques for graph analysis and parallel programming, and the practical application of data science for such tasks as building recommender systems or performing sentiment analysis. Topics and features: provides numerous practical case studies using real-world data throughout the book; supports understanding through hands-on experience of solving data science problems using Python; describes techniques and tools for statistical analysis, machine learning, graph analysis, and parallel programming; reviews a range of applications of data science, including recommender systems and sentiment analysis of text data; provides supplementary code resources and data at an associated website.

hands on data science with sql server 2017: Computational Science and Its Applications - ICCSA 2017 Osvaldo Gervasi, Beniamino Murgante, Sanjay Misra, Giuseppe Borruso, Carmelo M. Torre, Ana Maria A.C. Rocha, David Taniar, Bernady O. Apduhan, Elena Stankova, Alfredo Cuzzocrea, 2017-07-13 The six-volume set LNCS 10404-10409 constitutes the refereed proceedings of the 17th International Conference on Computational Science and Its Applications, ICCSA 2017, held in Trieste, Italy, in July 2017. The 313 full papers and 12 short papers included in the 6-volume proceedings set were carefully reviewed and selected from 1052 submissions. Apart from the general tracks, ICCSA 2017 included 43 international workshops in various areas of computational sciences, ranging from computational science technologies to specific areas of computational sciences, such as computer graphics and virtual reality. Furthermore, this year ICCSA 2017 hosted the XIV International Workshop On Quantum Reactive Scattering. The program also featured 3 keynote speeches and 4 tutorials.

hands on data science with sql server 2017: Big Data Analytics with Hadoop 3 Sridhar Alla, 2018-05-31 Explore big data concepts, platforms, analytics, and their applications using the power of Hadoop 3 Key Features Learn Hadoop 3 to build effective big data analytics solutions on-premise and on cloud Integrate Hadoop with other big data tools such as R, Python, Apache Spark, and Apache Flink Exploit big data using Hadoop 3 with real-world examples Book Description Apache Hadoop is the most popular platform for big data processing, and can be combined with a host of other big data tools to build powerful analytics solutions. Big Data Analytics with Hadoop 3 shows you how to do just that, by providing insights into the software as well as its benefits with the help of practical examples. Once you have taken a tour of Hadoop 3's latest features, you will get an overview of HDFS, MapReduce, and YARN, and how they enable faster, more efficient big data processing. You will then move on to learning how to integrate Hadoop with the open source tools, such as Python and R, to analyze and visualize data and perform statistical computing on big data. As you get acquainted with all this, you will explore how to use Hadoop 3 with Apache Spark and Apache Flink for real-time data analytics and stream processing. In addition to this, you will understand how to use Hadoop to build analytics solutions on the cloud and an end-to-end pipeline to perform big data analysis using practical use cases. By the end of this book, you will be well-versed with the analytical capabilities of the Hadoop ecosystem. You will be able to build powerful solutions to perform big data analytics and get insight effortlessly. What you will learn Explore the new features of Hadoop 3 along with HDFS, YARN, and MapReduce Get well-versed with the analytical capabilities of Hadoop ecosystem using practical examples Integrate Hadoop with R and Python for more efficient big data processing Learn to use Hadoop with Apache Spark and Apache Flink for real-time data analytics Set up a Hadoop cluster on AWS cloud Perform big data analytics on AWS using Elastic Map Reduce Who this book is for Big Data Analytics with Hadoop 3 is

for you if you are looking to build high-performance analytics solutions for your enterprise or business using Hadoop 3's powerful features, or you're new to big data analytics. A basic understanding of the Java programming language is required.

hands on data science with sql server 2017: *Joe Celko's Thinking in Sets: Auxiliary, Temporal, and Virtual Tables in SQL* Joe Celko, 2008-01-22 Perfectly intelligent programmers often struggle when forced to work with SQL. Why? Joe Celko believes the problem lies with their procedural programming mindset, which keeps them from taking full advantage of the power of declarative languages. The result is overly complex and inefficient code, not to mention lost productivity. This book will change the way you think about the problems you solve with SQL programs.. Focusing on three key table-based techniques, Celko reveals their power through detailed examples and clear explanations. As you master these techniques, you'll find you are able to conceptualize problems as rooted in sets and solvable through declarative programming. Before long, you'll be coding more quickly, writing more efficient code, and applying the full power of SQL - Filled with the insights of one of the world's leading SQL authorities - noted for his knowledge and his ability to teach what he knows - Focuses on auxiliary tables (for computing functions and other values by joins), temporal tables (for temporal queries, historical data, and audit information), and virtual tables (for improved performance) - Presents clear guidance for selecting and correctly applying the right table technique

hands on data science with sql server 2017: Data Science with SQL Server Quick Start Guide Dejan Sarka, 2018-08-31 Get unique insights from your data by combining the power of SQL Server, R and Python Key Features Use the features of SQL Server 2017 to implement the data science project life cycle Leverage the power of R and Python to design and develop efficient data models find unique insights from your data with powerful techniques for data preprocessing and analysis Book Description SQL Server only started to fully support data science with its two most recent editions. If you are a professional from both worlds, SQL Server and data science, and interested in using SQL Server and Machine Learning (ML) Services for your projects, then this is the ideal book for you. This book is the ideal introduction to data science with Microsoft SQL Server and In-Database ML Services. It covers all stages of a data science project, from business and data understanding, through data overview, data preparation, modeling and using algorithms, model evaluation, and deployment. You will learn to use the engines and languages that come with SQL Server, including ML Services with R and Python languages and Transact-SQL. You will also learn how to choose which algorithm to use for which task, and learn the working of each algorithm. What you will learn Use the popular programming languages, T-SQL, R, and Python, for data science Understand your data with queries and introductory statistics Create and enhance the datasets for ML Visualize and analyze data using basic and advanced graphs Explore ML using unsupervised and supervised models Deploy models in SQL Server and perform predictions Who this book is for SQL Server professionals who want to start with data science, and data scientists who would like to start using SQL Server in their projects will find this book to be useful. Prior exposure to SQL Server will be helpful.

hands on data science with sql server 2017: *SQL Server 2016 Developer's Guide* Dejan Sarka, Milos Radivojevic, William Durkin, 2017-03-22 Get the most out of the rich development capabilities of SQL Server 2016 to build efficient database applications for your organization About This Book Utilize the new enhancements in Transact-SQL and security features in SQL Server 2016 to build efficient database applications Work with temporal tables to get information about data stored in the table at any point in time A detailed guide to SQL Server 2016, introducing you to multiple new features and enhancements to improve your overall development experience Who This Book Is For This book is for database developers and solution architects who plan to use the new SQL Server 2016 features for developing efficient database applications. It is also ideal for experienced SQL Server developers who want to switch to SQL Server 2016 for its rich development capabilities. Some understanding of the basic database concepts and Transact-SQL language is assumed. What You Will Learn Explore the new development features introduced in SQL Server

2016 Identify opportunities for In-Memory OLTP technology, significantly enhanced in SQL Server 2016 Use columnstore indexes to get significant storage and performance improvements Extend database design solutions using temporal tables Exchange JSON data between applications and SQL Server in a more efficient way Migrate historical data transparently and securely to Microsoft Azure by using Stretch Database Use the new security features to encrypt or to have more granular control over access to rows in a table Simplify performance troubleshooting with Query Store Discover the potential of R's integration with SQL Server In Detail Microsoft SQL Server 2016 is considered the biggest leap in the data platform history of the Microsoft, in the ongoing era of Big Data and data science. Compared to its predecessors, SQL Server 2016 offers developers a unique opportunity to leverage the advanced features and build applications that are robust, scalable, and easy to administer. This book introduces you to new features of SQL Server 2016 which will open a completely new set of possibilities for you as a developer. It prepares you for the more advanced topics by starting with a quick introduction to SQL Server 2016's new features and a recapitulation of the possibilities you may have already explored with previous versions of SQL Server. The next part introduces you to small delights in the Transact-SQL language and then switches to a completely new technology inside SQL Server - JSON support. We also take a look at the Stretch database, security enhancements, and temporal tables. The last chapters concentrate on implementing advanced topics, including Query Store, columnstore indexes, and In-Memory OLTP. You will finally be introduced to R and how to use the R language with Transact-SQL for data exploration and analysis. By the end of this book, you will have the required information to design efficient, high-performance database applications without any hassle. Style and approach This book is a detailed guide to mastering the development features offered by SQL Server 2016, with a unique learn-as-you-do approach. All the concepts are explained in a very easy-to-understand manner and are supplemented with examples to ensure that you—the developer—are able to take that next step in building more powerful, robust applications for your organization with ease.

hands on data science with sql server 2017: Machine Learning for Hackers Drew Conway, John Myles White, 2012-02-13 If you're an experienced programmer interested in crunching data, this book will get you started with machine learning—a toolkit of algorithms that enables computers to train themselves to automate useful tasks. Authors Drew Conway and John Myles White help you understand machine learning and statistics tools through a series of hands-on case studies, instead of a traditional math-heavy presentation. Each chapter focuses on a specific problem in machine learning, such as classification, prediction, optimization, and recommendation. Using the R programming language, you'll learn how to analyze sample datasets and write simple machine learning algorithms. Machine Learning for Hackers is ideal for programmers from any background, including business, government, and academic research. Develop a naïve Bayesian classifier to determine if an email is spam, based only on its text Use linear regression to predict the number of page views for the top 1,000 websites Learn optimization techniques by attempting to break a simple letter cipher Compare and contrast U.S. Senators statistically, based on their voting records Build a “whom to follow” recommendation system from Twitter data

hands on data science with sql server 2017: Practical SQL, 2nd Edition Anthony DeBarros, 2022-01-25 Analyze data like a pro, even if you're a beginner. 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. Anthony DeBarros, a journalist and data analyst, focuses on using SQL to find the story within your data. The examples and code use the open-source database PostgreSQL and its companion pgAdmin interface, and the concepts you learn will apply to most database management systems, including MySQL, Oracle, SQLite, and others.* You'll first cover the fundamentals of databases and the SQL language, then build skills by analyzing data from real-world datasets such as US Census demographics, New York City taxi rides, and earthquakes from US Geological Survey. Each chapter includes exercises and examples that 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 Aggregate, sort, and filter data to find patterns Use functions for basic math and advanced statistical operations Identify errors in data and clean them up Analyze spatial data with a geographic information system (PostGIS) Create advanced queries and automate tasks This updated second edition has been thoroughly revised to reflect the latest in SQL features, including additional advanced query techniques for wrangling data. This edition also has two new chapters: an expanded set of instructions on for setting up your system plus a chapter on using PostgreSQL with the popular JSON data interchange format. 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. * Microsoft SQL Server employs a variant of the language called T-SQL, which is not covered by Practical SQL.

Back to Home: <https://a.comtex-nj.com>