

## Apache Spark Book Soup

Summary Hello! Python fully covers the building blocks of Python programming and gives you a gentle introduction to more advanced topics such as object-oriented programming, functional programming, network programming, and program design. New (or nearly new) programmers will learn most of what they need to know to start using Python immediately. About this Book Programmers love Python because it's fast and efficient. Shouldn't learning Python be just the same? Hello! Python starts quickly and simply, with a line of Python code. You'll learn the basics the right way--by writing your own programs. Along the way, you'll get a gentle introduction to more advanced concepts and new programming styles.> No experience with Python needed. Exposure to another programming language is helpful but not required. Purchase of the print book comes with an offer of a free PDF, ePub, and Kindle eBook from Manning. Also available is all code from the book. What Makes Hello! Python special Learn Python fast Even if you've never written a line of code before, you'll be writing real Python apps in just an hour or two. Great examples There's something new in every chapter, including games, web programming with Django, databases, and more. User Friendly guides Using lots of illustrations and a down-to-earth writing style, this book invites you to explore Python along with half-a-dozen traveling companions from the User Friendly cartoon strip.

=====  
?== Table of Contents Why Python? Hunt the Wumpus Interacting with theWorld Getting Organized Business-Oriented Programming Classes and Object-oriented Programming Sufficiently Advanced Technology Django! Gaming with Pyglet Twisted Networking Django Revisted! Where to from Here?

## Read Online Apache Spark Book Soup

Develop applications for the big data landscape with Spark and Hadoop. This book also explains the role of Spark in developing scalable machine learning and analytics applications with Cloud technologies. Beginning Apache Spark 2 gives you an introduction to Apache Spark and shows you how to work with it. Along the way, you'll discover resilient distributed datasets (RDDs); use Spark SQL for structured data; and learn stream processing and build real-time applications with Spark Structured Streaming. Furthermore, you'll learn the fundamentals of Spark ML for machine learning and much more. After you read this book, you will have the fundamentals to become proficient in using Apache Spark and know when and how to apply it to your big data applications. What You Will Learn Understand Spark unified data processing platform How to run Spark in Spark Shell or Databricks Use and manipulate RDDs Deal with structured data using Spark SQL through its operations and advanced functions Build real-time applications using Spark Structured Streaming Develop intelligent applications with the Spark Machine Learning library Who This Book Is For Programmers and developers active in big data, Hadoop, and Java but who are new to the Apache Spark platform.

Open the pages of this lively book and discover a bowl of animal alphabet soup sure to spark children's imaginations. From the tiny ant on the front endpapers to the exuberant zebra in the final spread, children will delight in the mischievous menagerie that bursts onto the table, while at the same time learning their letters and animals. Filled with imaginative details, it is a visual feast from start to finish.

Historically, nursing, in all of its missions of research/scholarship, education and practice, has not had access to large patient databases. Nursing consequently adopted qualitative methodologies with small sample sizes, clinical trials and lab research. Historically, large data

## Read Online Apache Spark Book Soup

methods were limited to traditional biostatistical analyses. In the United States, large payer data has been amassed and structures/organizations have been created to welcome scientists to explore these large data to advance knowledge discovery. Health systems electronic health records (EHRs) have now matured to generate massive databases with longitudinal trending. This text reflects how the learning health system infrastructure is maturing, and being advanced by health information exchanges (HIEs) with multiple organizations blending their data, or enabling distributed computing. It educates the readers on the evolution of knowledge discovery methods that span qualitative as well as quantitative data mining, including the expanse of data visualization capacities, are enabling sophisticated discovery. New opportunities for nursing and call for new skills in research methodologies are being further enabled by new partnerships spanning all sectors.

Do you have a real relationship with God, or do you just have a religion? Do you know God, or do you just know about God? In *How Big Is Your God?* Paul Coutinho, SJ, challenges us to grow stronger and deeper in our faith and in our relationship with God—a God whose love knows no bounds. To help us on our way, Coutinho introduces us to people in various world religions—from Hindu friends to Buddhist teachers to St. Ignatius of Loyola—who have shaped his spiritual life and made possible his deep, personal relationship with God.

The current political climate has left many of us wondering how our government actually operates. Sure, we learned about it in school, but if put to the test, how many of us can correctly explain the branches of government? The history of politics? The differences and connections between local government and federal government? Enter *How to Be an American*. While author and illustrator Silvia Hidalgo was studying for her citizenship test, she

## Read Online Apache Spark Book Soup

quickly found that the materials provided by the government were lacking. In order to more easily absorb the information, Hidalgo started her own illustrated reference to civics facts and American history essentials. She's collected her findings in *How to Be an American*, a freshly designed and illustrated two-color guide to all things America.

Build, process and analyze large-scale graph data effectively with Spark About This Book Find solutions for every stage of data processing from loading and transforming graph data to Improve the scalability of your graphs with a variety of real-world applications with complete Scala code. A concise guide to processing large-scale networks with Apache Spark. Who This Book Is For This book is for data scientists and big data developers who want to learn the processing and analyzing graph datasets at scale. Basic programming experience with Scala is assumed. Basic knowledge of Spark is assumed. What You Will Learn Write, build and deploy Spark applications with the Scala Build Tool. Build and analyze large-scale network datasets Analyze and transform graphs using RDD and graph-specific operations Implement new custom graph operations tailored to specific needs. Develop iterative and efficient graph algorithms using message aggregation and Pregel abstraction Extract subgraphs and use it to discover common clusters Analyze graph data and solve various data science problems using real-world datasets. In Detail Apache Spark is the next standard of open-source cluster-computing engine for processing big data. Many practical computing problems concern large graphs, like the Web graph and various social networks. The scale of these graphs - in some cases billions of vertices, trillions of edges - poses challenges to their efficient processing. Apache Spark GraphX API combines the advantages of both data-parallel and graph-parallel systems by efficiently expressing graph computation within the Spark data-parallel framework.

## Read Online Apache Spark Book Soup

This book will teach the user to do graphical programming in Apache Spark, apart from an explanation of the entire process of graphical data analysis. You will journey through the creation of graphs, its uses, its exploration and analysis and finally will also cover the conversion of graph elements into graph structures. This book begins with an introduction of the Spark system, its libraries and the Scala Build Tool. Using a hands-on approach, this book will quickly teach you how to install and leverage Spark interactively on the command line and in a standalone Scala program. Then, it presents all the methods for building Spark graphs using illustrative network datasets. Next, it will walk you through the process of exploring, visualizing and analyzing different network characteristics. This book will also teach you how to transform raw datasets into a usable form. In addition, you will learn powerful operations that can be used to transform graph elements and graph structures. Furthermore, this book also teaches how to create custom graph operations that are tailored for specific needs with efficiency in mind. The later chapters of this book cover more advanced topics such as clustering graphs, implementing graph-parallel iterative algorithms and learning methods from graph data. Style and approach A step-by-step guide that will walk you through the key ideas and techniques for processing big graph data at scale, with practical examples that will ensure an overall understanding of the concepts of Spark.

Every team needs a leader, but why do we so often take that to mean that the appropriate workplace team needs to consist of one gem of a worker complemented with a bunch of obedient order-takers and yes men? What if the complementary fits between the team members were not with how well they performed the tasks handed down to them but with how they all used their unique strengths to share knowledge, push the envelope, and lead together

## Read Online Apache Spark Book Soup

in the challenge before them?The team of authors behind *A Team of Leaders* wants to show readers how to design systems within their organization and management procedures that nurture the leadership potential of every employee, not just the ones they ear-marked as having potential for promotion. The proven principles and techniques within these invaluable pages include:

- The Five-Stage Team Development Model that maps the transition from traditional to self-directed teams
- Best practices in team process design
- A Team Value Creation Tool that allows members to appreciate the significance of what they contribute each day
- Visual Management
- And more

The key to your company's success is creating successful teams of leaders combining their individual talents and strengths into a single, unstoppable driving force. The fresh approach taught in this indispensable guide will transform passive groups of disparate people into the effective teams of leaders you didn't know was possible to have.

Reveals how the voyages of Columbus reintroduced plants and animals that had been separated millions of years earlier, documenting how the ensuing exchange of flora and fauna between Eurasia and the Americas fostered a European rise, decimated imperial China and rendered Manila and Mexico City the center of the world for two centuries. Reprint.

The process of user-centered innovation: how it can benefit both users and manufacturers and how its emergence will bring changes in business models and in public policy. Innovation is rapidly becoming democratized. Users, aided by improvements in computer and communications technology, increasingly can develop their own new products and services. These innovating users—both

individuals and firms—often freely share their innovations with others, creating user-innovation communities and a rich intellectual commons. In *Democratizing Innovation*, Eric von Hippel looks closely at this emerging system of user-centered innovation. He explains why and when users find it profitable to develop new products and services for themselves, and why it often pays users to reveal their innovations freely for the use of all. The trend toward democratized innovation can be seen in software and information products—most notably in the free and open-source software movement—but also in physical products. Von Hippel's many examples of user innovation in action range from surgical equipment to surfboards to software security features. He shows that product and service development is concentrated among "lead users," who are ahead on marketplace trends and whose innovations are often commercially attractive. Von Hippel argues that manufacturers should redesign their innovation processes and that they should systematically seek out innovations developed by users. He points to businesses—the custom semiconductor industry is one example—that have learned to assist user-innovators by providing them with toolkits for developing new products. User innovation has a positive impact on social welfare, and von Hippel proposes that government policies, including R&D subsidies and tax credits, should be realigned to eliminate biases against it. The

## Read Online Apache Spark Book Soup

goal of a democratized user-centered innovation system, says von Hippel, is well worth striving for. An electronic version of this book is available under a Creative Commons license.

Red Sparrow is now a major motion picture starring Jennifer Lawrence and Joel Edgerton! The thrilling sequel to Red Sparrow—CIA insider Jason Matthews’s compulsively readable New York Times bestseller and Edgar Award winner—featuring Russian spy Dominika Egorova and CIA agent Nate Nash “shimmers with authenticity. The villains are richly drawn...the scenes of them on the job are beyond chilling” (The New York Times Book Review). Captain Dominika Egorova of the Russian Intelligence Service despises the oligarchs, crooks, and thugs of Putin’s Russia—but what no one knows is that she is also working for the CIA. Her “sparrow” training in the art of sexual espionage further complicates the mortal risks she must take, as does her love for her handler Nate Nash—a shared lust that is as dangerous as treason. As Dominika expertly dodges exposure, she deals with a murderously psychotic boss, survives an Iranian assassination attempt and attempts to rescue an arrested double agent—and thwart Putin’s threatening flirtations. A grand, wildly entertaining ride through the steel-trap mind of a CIA insider, Palace of Treason is a story “as suspenseful and cinematic as the best spy movies” (The Philadelphia

## Read Online Apache Spark Book Soup

Inquirer)—one that feels fresh and so possible, in fact, that it's doubtful this novel can ever be published in Russia.

“Atkinson and Jewell invite each of us to reimagine one's connection to the land while cultivating nature close to home. A must-read for anyone searching for inspired solutions for designing or refining a garden.” —founder of Pass the Pistil From windswept deserts to misty seaside hills and verdant valleys, the natural landscapes of the American West offer an astounding variety of climates for gardens. *Under Western Skies* reveals thirty-six of the most innovative designs—all embracing and celebrating the very soul of the land on which they grow. For the gardeners featured here, nature is the ultimate inspiration rather than something to be dominated, and *Under Western Skies* shows the strong connection each garden has with its place. Packed with Atkinson's stunning photographs and illuminated by Jewell's deep interest in the relationships between people and the spaces they inhabit, *Under Western Skies* offers page after page of encouraging ingenuity and inventive design for passionate gardeners who call the West home.

This latest edition of an official U.S. Government military history classic provides an authoritative historical survey of the organization and accomplishments of the United States Army. This scholarly yet readable book is designed to inculcate an

## Read Online Apache Spark Book Soup

awareness of our nation's military past and to demonstrate that the study of military history is an essential ingredient in leadership development. It is also an essential addition to any personal military history library. This text is used in military ROTC training courses as a basic military history textbook. Volume 1 of 2 volume set.

California Friendly® is California's future. Water reliability is dependent on using water wisely. We need to create sustainable gardens that rely on less water. This maintenance guide will help you support California's future: \*Uncover the secrets of efficient irrigation. \*Explore the techniques for irrigating with recycled water. \*Get the maintenance tips for hundreds of California Friendly® plants. \*Discover the methods and means of managing weed and pest infestations. \*Learn how to maintain rainwater capture opportunities. This book has been written for every landscaper, gardener and land manager in Southern California. It has been produced by the very first collaboration between three Southern California organizations, LADWP, MWD and SoCalGas. Grab a copy--they are free--use the information in your garden and help us create a beautiful, thriving and sustainable future.

In response to a recent surge of interest in Native American history, culture, and lore, Hippocrene brings you a concise and straightforward dictionary of the

## Read Online Apache Spark Book Soup

Navajo tongue. The dictionary is designed to aid Navajos learning English as well as English speakers interested in acquiring knowledge of Navajo. The largest of all the Native American tribes, the Navajo number about 125,000 and live mostly on reservations in Arizona, New Mexico, and Utah. Over 9,000 entries; A detailed section on Navajo pronunciation; A comprehensive, modern vocabulary; Useful, everyday expressions.

This book provides a complete and modern guide to web scraping, using Python as the programming language, without glossing over important details or best practices. Written with a data science audience in mind, the book explores both scraping and the larger context of web technologies in which it operates, to ensure full understanding. The authors recommend web scraping as a powerful tool for any data scientist's arsenal, as many data science projects start by obtaining an appropriate data set. Starting with a brief overview on scraping and real-life use cases, the authors explore the core concepts of HTTP, HTML, and CSS to provide a solid foundation. Along with a quick Python primer, they cover Selenium for JavaScript-heavy sites, and web crawling in detail. The book finishes with a recap of best practices and a collection of examples that bring together everything you've learned and illustrate various data science use cases. What You'll Learn Leverage well-established best practices and commonly-used

## Read Online Apache Spark Book Soup

Python packages Handle today's web, including JavaScript, cookies, and common web scraping mitigation techniques Understand the managerial and legal concerns regarding web scraping Who This Book is For A data science oriented audience that is probably already familiar with Python or another programming language or analytical toolkit (R, SAS, SPSS, etc). Students or instructors in university courses may also benefit. Readers unfamiliar with Python will appreciate a quick Python primer in chapter 1 to catch up with the basics and provide pointers to other guides as well.

A sampling of the writings of Chinese women poets from ancient times to the twentieth century is accompanied by brief biographies and a discussion of their influence on Chinese literature.

### Getting Started with Beautiful SoupPackt Publishing Ltd

With Michael missing, the members of the Electrobian are divided, leaving them vulnerable to Hatch's forces.

Summary Natural Language Processing in Action is your guide to creating machines that understand human language using the power of Python with its ecosystem of packages dedicated to NLP and AI. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Recent advances in deep learning empower applications to understand text and speech with extreme accuracy. The result? Chatbots that can imitate real people, meaningful resume-to-job matches, superb

## Read Online Apache Spark Book Soup

predictive search, and automatically generated document summaries—all at a low cost. New techniques, along with accessible tools like Keras and TensorFlow, make professional-quality NLP easier than ever before. About the Book Natural Language Processing in Action is your guide to building machines that can read and interpret human language. In it, you'll use readily available Python packages to capture the meaning in text and react accordingly. The book expands traditional NLP approaches to include neural networks, modern deep learning algorithms, and generative techniques as you tackle real-world problems like extracting dates and names, composing text, and answering free-form questions. What's inside Some sentences in this book were written by NLP! Can you guess which ones? Working with Keras, TensorFlow, gensim, and scikit-learn Rule-based and data-based NLP Scalable pipelines About the Reader This book requires a basic understanding of deep learning and intermediate Python skills. About the Author Hobson Lane, Cole Howard, and Hannes Max Hapke are experienced NLP engineers who use these techniques in production. Table of Contents PART 1 - WORDY MACHINES Packets of thought (NLP overview) Build your vocabulary (word tokenization) Math with words (TF-IDF vectors) Finding meaning in word counts (semantic analysis) PART 2 - DEEPER LEARNING (NEURAL NETWORKS) Baby steps with neural networks (perceptrons and backpropagation) Reasoning with word vectors (Word2vec) Getting words in order with convolutional neural networks (CNNs) Loopy (recurrent) neural networks (RNNs) Improving retention with long short-term memory networks Sequence-to-sequence models and attention PART 3 - GETTING REAL (REAL-WORLD NLP CHALLENGES) Information extraction (named entity extraction and question answering) Getting chatty (dialog engines) Scaling up (optimization, parallelization, and batch processing)

## Read Online Apache Spark Book Soup

Explores the homogenization of American culture and the impact of the fast food industry on modern-day health, economy, politics, popular culture, entertainment, and food production. Work with petabyte-scale datasets while building a collaborative, agile workplace in the process. This practical book is the canonical reference to Google BigQuery, the query engine that lets you conduct interactive analysis of large datasets. BigQuery enables enterprises to efficiently store, query, ingest, and learn from their data in a convenient framework. With this book, you'll examine how to analyze data at scale to derive insights from large datasets efficiently. Valliappa Lakshmanan, tech lead for Google Cloud Platform, and Jordan Tigani, engineering director for the BigQuery team, provide best practices for modern data warehousing within an autoscaled, serverless public cloud. Whether you want to explore parts of BigQuery you're not familiar with or prefer to focus on specific tasks, this reference is indispensable.

This book is a practical, hands-on guide that takes you through the techniques of web scraping using Beautiful Soup. Getting Started with Beautiful Soup is great for anybody who is interested in website scraping and extracting information. However, a basic knowledge of Python, HTML tags, and CSS is required for better understanding.

Playing in puddles is an enjoyable experience that children love to do! Splashing in the water, building puddle boats, and using the water to make 'puddle soup' are all favourite activities that also integrate rich learning. Heading outside to play on a wet, rainy day uses children's excitement and enthusiasm for sensory experiences to inspire rich mathematical and scientific inquiries about what they are seeing and experiencing. In this book readers explore the math that exists in puddle play using the 'see, think, and wonder' thinking routine. Colourful photos

## Read Online Apache Spark Book Soup

and an easy to access text invite both children and adults to explore the fun and excitement of playing with water outdoors. This book can spark mathematical conversations with children, and be used as a guide for discovering the rich math that exists in weather-based activities. Photos can also be used without text to engage children in math talks as they observe and discuss what they see on each page. Educators will find this book to be an effective resource to inspire math talks within the classroom or before heading outside to play. Often educators are asked to justify messy outdoor play in the puddles. Exploring Puddle Math: See, Think and Wonder before heading outside will help empower children to notice and name the learning they are experiencing while having fun in the water. Families can use this book as a tool for helping inspire math thinking and conversation with children during quality reading time together or before heading outside to play.

This is a book about the big questions in life: knowledge, consciousness, fate, God, truth, goodness, justice. It is for anyone who believes there are big questions out there, but does not know how to approach them. Think sets out to explain what they are and why they are important. Simon Blackburn begins by putting forward a convincing case for the study of philosophy and goes on to give the reader a sense of how the great historical figures such as Descartes, Hume, Kant, and Wittgenstein have approached its central themes. Each chapter explains a major issue, and gives the reader a self-contained guide through the problems that philosophers have studied. The large scope of topics covered range from scepticism, the self, mind and body, and freedom to ethics and the arguments surrounding the existence of God. Lively and approachable, this book is ideal for all those who want to learn how the basic techniques of thinking shape our existence.

## Read Online Apache Spark Book Soup

Whether Jeffrey McDaniel is denouncing insomnia ("4,000 A.M."), exploring family tragedy ("Ghost Townhouse"), or celebrating love and lust ("The Biology of Numbers"), his writing is original and provocative. A noted poet, McDaniel has appeared on ABC's Nightline and NPR's Talk of the Nation. "Wild, fierce, irreverent, full of praise and lament, and deeply, intensely human." — Thomas Lux

Hungry? Got a craving for some ooey, gooey slime? Learn how to master the art of culinary slime making with YouTuber and slime Chef, Krist Soup. Included is a full menu of over 30 DIY slime meal recipes for Breakfast, Lunch, Dinner, and Dessert. Each recipe comes with step by step instructions and guided images to help relieve stress, enhance creativity, and bring some serious satisfaction to the brain and soul. **MADE WITH LOVE AND ALL NATURAL GLUE.**

"Slime Soup" covers how to make basic slime and various types of some of the most popular slime. (Jiggly Slime, Butter Slime, Crunchy Slime, Fluffy Slime, etc.) "I created this book to spark creativity in all individuals. Not only do I want people to have fun while making slime, but also for people to get interested in learning how to cook meals. Plus, it's the perfect excuse to play with your food."

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;

## Read Online Apache Spark Book Soup

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.

When Grace's parents and siblings are murdered by the Guiltless Gang for their Arizona horse ranch outside Tombstone, she vows to devote her life to revenge--but the Chiricahua she finds sanctuary with try to teach her a better way.

Uses picture books to teach such key scientific topics as seeds and plants, habitats, light and color, and water cycle and includes science information, management tips, book connections, Web sites, and assessment ideas.

Completion of the first phase of the Human Genome Project has presented scientists with a mountain of new information. The availability of all human genes and their locations is exciting, but their mechanisms of action and interaction with other genes are often unknown. Certain variations in the environment and characteristics of human genes make it difficult for scientists to maximize their efforts to take advantage of this new fund of knowledge. The high degree of similarity between mouse and human genomes and physiology has made mice the laboratory animals of choice in the search for models of human disease. In recent years significant advances in our ability to examine the mouse eye and to view vascular and neural structures in vivo has led to a marked increase in interest in mouse ocular pathology and development. Besides providing much new information, *Systematic Evaluation of the Mouse Eye: Anatomy, Pathology, and Biomethods* provides investigators with a resource that summarizes the widely scattered information on mouse eye anatomy, development, and pathology. It describes the

## Read Online Apache Spark Book Soup

normal anatomy, pre- and postnatal development, regional pathology, and methodology for evaluation of the mouse eye and adnexae. The first section covers the principal features of mouse eye anatomy and development. The second part discusses issues in evaluation of mutations in mice, including background ocular disease, control selection, quantitative traits, and ENU mutagenesis. With this background in place, the book then reviews and illustrates regional ocular pathology and correlates it with human eye disease. Many of the contributors who have helped develop and refine the specialized techniques necessary for the examination and preparation of mouse eye tissue for microscopy provide "how-to" chapters. These chapters cover methods for necropsy, photography, histopathology, confocal microscopy, intraocular pressure measurement, and electroretinography. Systematic Evaluation of the Mouse Eye: Anatomy, Pathology, and Biomethods provides the first comprehensive and systematic guide to understanding the development, structure, and abnormal pathology of the mouse eye so you can make the most of your genetic research.

The Visual Miscellaneum is a unique, groundbreaking look at the modern information age, helping readers make sense of the countless statistics and random facts that constantly bombard us. Using cutting edge graphs, charts, and illustrations, David McCandless creatively visualizes the world's surprising relationships and compelling data, covering everything from the most pleasurable guilty pleasures to how long it takes different condiments to spoil to world maps of Internet search terms.

How do you take your data analysis skills beyond Excel to the next level? By learning just enough Python to get stuff done. This hands-on guide shows non-programmers like you how to process information that's initially too messy or difficult to access. You don't need to know a

## Read Online Apache Spark Book Soup

thing about the Python programming language to get started. Through various step-by-step exercises, you'll learn how to acquire, clean, analyze, and present data efficiently. You'll also discover how to automate your data process, schedule file- editing and clean-up tasks, process larger datasets, and create compelling stories with data you obtain. Quickly learn basic Python syntax, data types, and language concepts Work with both machine-readable and human-consumable data Scrape websites and APIs to find a bounty of useful information Clean and format data to eliminate duplicates and errors in your datasets Learn when to standardize data and when to test and script data cleanup Explore and analyze your datasets with new Python libraries and techniques Use Python solutions to automate your entire data-wrangling process Get Started Fast with Apache Hadoop® 2, YARN, and Today's Hadoop Ecosystem With Hadoop 2.x and YARN, Hadoop moves beyond MapReduce to become practical for virtually any type of data processing. Hadoop 2.x and the Data Lake concept represent a radical shift away from conventional approaches to data usage and storage. Hadoop 2.x installations offer unmatched scalability and breakthrough extensibility that supports new and existing Big Data analytics processing methods and models. Hadoop® 2 Quick-Start Guide is the first easy, accessible guide to Apache Hadoop 2.x, YARN, and the modern Hadoop ecosystem. Building on his unsurpassed experience teaching Hadoop and Big Data, author Douglas Eadline covers all the basics you need to know to install and use Hadoop 2 on personal computers or servers, and to navigate the powerful technologies that complement it. Eadline concisely introduces and explains every key Hadoop 2 concept, tool, and service, illustrating each with a simple "beginning-to-end" example and identifying trustworthy, up-to-date resources for learning more. This guide is ideal if you want to learn about Hadoop 2 without getting mired in technical

## Read Online Apache Spark Book Soup

details. Douglas Eadline will bring you up to speed quickly, whether you're a user, admin, devops specialist, programmer, architect, analyst, or data scientist. Coverage Includes Understanding what Hadoop 2 and YARN do, and how they improve on Hadoop 1 with MapReduce Understanding Hadoop-based Data Lakes versus RDBMS Data Warehouses Installing Hadoop 2 and core services on Linux machines, virtualized sandboxes, or clusters Exploring the Hadoop Distributed File System (HDFS) Understanding the essentials of MapReduce and YARN application programming Simplifying programming and data movement with Apache Pig, Hive, Sqoop, Flume, Oozie, and HBase Observing application progress, controlling jobs, and managing workflows Managing Hadoop efficiently with Apache Ambari—including recipes for HDFS to NFSv3 gateway, HDFS snapshots, and YARN configuration Learning basic Hadoop 2 troubleshooting, and installing Apache Hue and Apache Spark

This book describes the paleomagnetism of sediments and sedimentary rocks, how sediments and sedimentary rocks become magnetized, and how the physical and chemical processes involved can affect the accuracy of paleomagnetism. Topics covered include depositional and post-depositional remanence acquisition, the detection and correction of compaction-caused inclination shallowing, reduction diagenesis of magnetic minerals, chemical remagnetization, and rotation of remanence by grain-scale rock strain. The book also has a chapter on environmental paleomagnetism, including examples of the new technique of high-resolution rock magnetic cyclostratigraphy and its application to sedimentary sequences. By emphasizing the accuracy of sedimentary paleomagnetism and the magnitude of post-depositional processes that can affect it, the book will be invaluable in the geologic interpretation of

## Read Online Apache Spark Book Soup

sedimentary paleomagnetic data. Paleomagnetism of Sedimentary Rocks will be welcomed by paleomagnetists, students of paleomagnetism and all Earth scientists who use sedimentary paleomagnetic data in their research. Additional resources for this book can be found at: [www.wiley.com/go/kodama/paleomagnetism](http://www.wiley.com/go/kodama/paleomagnetism).

The award-winning creator of *Blue's Clues*, *Super Why!*, and *Daniel Tiger's Neighborhood* shares the secret sauce behind her shows' powerful, transformative results in the form of eleven research-based, foundational "clues" to ensure that preschoolers flourish academically, socially, and emotionally during this critical time. The preschool years—when children are between the ages of two and five—are the most influential, important years in a child's life. Studies show that pausing to interact, playing to solve problems, diffusing with humor, and using repetition are the hidden clues conscious parents use to raise successful kids and help them learn critical thinking skills, foster empathy, and nurture their sense of self-worth. Angela C. Santomero, MA, the award-winning creator of children's television phenomena knows this better than anyone and has spent decades working to instill confidence in her young viewers. In *Preschool Clues*, she breaks down the philosophy behind her shows—educating, inspiring, and empowering kids—into concrete strategies that parents and educators can incorporate into their family and classroom to set their preschoolers up for success, such as:

- Intentionally pausing to foster bonding, independence, and resilience
- Developing empathy and confidence through soliciting preschoolers' help
- Becoming "fluent" in the language of preschoolers: Play
- Igniting your preschooler's curiosity
- Being an involved co-player everyday
- Designing a healthy media diet

In *Preschool Clues*, Angela shares the latest research from top thinkers in child development and education. Through her practical, straightforward advice and inspiring,

## Read Online Apache Spark Book Soup

conversational approach, you will not only understand exactly what your children are learning from the shows they watch and why these shows are so effective, you'll know exactly how to apply these same proven approaches in your daily life and with the same powerful results.

"The Book of This and That" by Robert Lynd. Published by Good Press. Good Press publishes a wide range of titles that encompasses every genre. From well-known classics & literary fiction and non-fiction to forgotten?or yet undiscovered gems?of world literature, we issue the books that need to be read. Each Good Press edition has been meticulously edited and formatted to boost readability for all e-readers and devices. Our goal is to produce eBooks that are user-friendly and accessible to everyone in a high-quality digital format.

Journalist Walls grew up with parents whose ideals and stubborn nonconformity were their curse and their salvation. Rex and Rose Mary and their four children lived like nomads, moving among Southwest desert towns, camping in the mountains. Rex was a charismatic, brilliant man who, when sober, captured his children's imagination, teaching them how to embrace life fearlessly. Rose Mary painted and wrote and couldn't stand the responsibility of providing for her family. When the money ran out, the Walls retreated to the dismal West Virginia mining town Rex had tried to escape. As the dysfunction escalated, the children had to fend for themselves, supporting one another as they found the resources and will to leave home. Yet Walls describes her parents with deep affection in this tale of unconditional love in a family that, despite its

## Read Online Apache Spark Book Soup

profound flaws, gave her the fiery determination to carve out a successful life. -- From publisher description.

Learn the art of efficient web scraping and crawling with Python About This Book Extract data from any source to perform real time analytics. Full of techniques and examples to help you crawl websites and extract data within hours. A hands-on guide to web scraping and crawling with real-life problems and solutions Who This Book Is For If you are a software developer, data scientist, NLP or machine-learning enthusiast or just need to migrate your company's wiki from a legacy platform, then this book is for you. It is perfect for someone , who needs instant access to large amounts of semi-structured data effortlessly. What You Will Learn Understand HTML pages and write XPath to extract the data you need Write Scrapy spiders with simple Python and do web crawls Push your data into any database, search engine or analytics system Configure your spider to download files, images and use proxies Create efficient pipelines that shape data in precisely the form you want Use Twisted Asynchronous API to process hundreds of items concurrently Make your crawler super-fast by learning how to tune Scrapy's performance Perform large scale distributed crawls with scrapyd and scrapinghub In Detail This book covers the long awaited Scrapy v 1.0 that empowers you to extract useful data from virtually any source with very little effort. It starts off by explaining the fundamentals of Scrapy framework, followed by a thorough description of how to extract data from any source, clean it up, shape it as per your requirement using

## Read Online Apache Spark Book Soup

Python and 3rd party APIs. Next you will be familiarised with the process of storing the scrapped data in databases as well as search engines and performing real time analytics on them with Spark Streaming. By the end of this book, you will perfect the art of scarping data for your applications with ease Style and approach It is a hands on guide, with first few chapters written as a tutorial, aiming to motivate you and get you started quickly. As the book progresses, more advanced features are explained with real world examples that can be reffered while developing your own web applications. Summary Modern data science solutions need to be clean, easy to read, and scalable. In Mastering Large Datasets with Python, author J.T. Wolohan teaches you how to take a small project and scale it up using a functionally influenced approach to Python coding. You'll explore methods and built-in Python tools that lend themselves to clarity and scalability, like the high-performing parallelism method, as well as distributed technologies that allow for high data throughput. The abundant hands-on exercises in this practical tutorial will lock in these essential skills for any large-scale data science project. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Programming techniques that work well on laptop-sized data can slow to a crawl—or fail altogether—when applied to massive files or distributed datasets. By mastering the powerful map and reduce paradigm, along with the Python-based tools that support it, you can write data-centric applications that scale efficiently without requiring codebase rewrites as your

## Read Online Apache Spark Book Soup

requirements change. About the book Mastering Large Datasets with Python teaches you to write code that can handle datasets of any size. You'll start with laptop-sized datasets that teach you to parallelize data analysis by breaking large tasks into smaller ones that can run simultaneously. You'll then scale those same programs to industrial-sized datasets on a cluster of cloud servers. With the map and reduce paradigm firmly in place, you'll explore tools like Hadoop and PySpark to efficiently process massive distributed datasets, speed up decision-making with machine learning, and simplify your data storage with AWS S3. What's inside An introduction to the map and reduce paradigm Parallelization with the multiprocessing module and pathos framework Hadoop and Spark for distributed computing Running AWS jobs to process large datasets About the reader For Python programmers who need to work faster with more data. About the author J. T. Wolohan is a lead data scientist at Booz Allen Hamilton, and a PhD researcher at Indiana University, Bloomington. Table of Contents: PART 1 1 | Introduction 2 | Accelerating large dataset work: Map and parallel computing 3 | Function pipelines for mapping complex transformations 4 | Processing large datasets with lazy workflows 5 | Accumulation operations with reduce 6 | Speeding up map and reduce with advanced parallelization PART 2 7 | Processing truly big datasets with Hadoop and Spark 8 | Best practices for large data with Apache Streaming and mrjob 9 | PageRank with map and reduce in PySpark 10 | Faster decision-making with machine learning and PySpark PART 3 11 | Large datasets in the cloud with Amazon Web

## Read Online Apache Spark Book Soup

Services and S3 12 | MapReduce in the cloud with Amazon's Elastic MapReduce Summary Introducing Data Science teaches you how to accomplish the fundamental tasks that occupy data scientists. Using the Python language and common Python libraries, you'll experience firsthand the challenges of dealing with data at scale and gain a solid foundation in data science. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Many companies need developers with data science skills to work on projects ranging from social media marketing to machine learning. Discovering what you need to learn to begin a career as a data scientist can seem bewildering. This book is designed to help you get started. About the Book Introducing Data Science Introducing Data Science explains vital data science concepts and teaches you how to accomplish the fundamental tasks that occupy data scientists. You'll explore data visualization, graph databases, the use of NoSQL, and the data science process. You'll use the Python language and common Python libraries as you experience firsthand the challenges of dealing with data at scale. Discover how Python allows you to gain insights from data sets so big that they need to be stored on multiple machines, or from data moving so quickly that no single machine can handle it. This book gives you hands-on experience with the most popular Python data science libraries, Scikit-learn and StatsModels. After reading this book, you'll have the solid foundation you need to start a career in data science. What's Inside Handling large data Introduction

## Read Online Apache Spark Book Soup

to machine learning Using Python to work with data Writing data science algorithms  
About the Reader This book assumes you're comfortable reading code in Python or a similar language, such as C, Ruby, or JavaScript. No prior experience with data science is required. About the Authors Davy Cielen, Arno D. B. Meysman, and Mohamed Ali are the founders and managing partners of Optimately and Maiton, where they focus on developing data science projects and solutions in various sectors. Table of Contents  
Data science in a big data world The data science process Machine learning Handling large data on a single computer First steps in big data Join the NoSQL movement The rise of graph databases Text mining and text analytics Data visualization to the end user

[Copyright: 03c62e07c84dd34b92195335c4462fbf](https://www.copyright.com/03c62e07c84dd34b92195335c4462fbf)