

---

# Read Book Pdf Free Action In Spark

---

This is likewise one of the factors by obtaining the soft documents of this **Pdf Free Action In Spark** by online. You might not require more become old to spend to go to the ebook creation as with ease as search for them. In some cases, you likewise reach not discover the revelation Pdf Free Action In Spark that you are looking for. It will certainly squander the time.

However below, once you visit this web page, it will be as a result utterly simple to acquire as without difficulty as download guide Pdf Free Action In Spark

It will not undertake many mature as we explain before. You can realize it even though fake something else at house and even in your workplace. fittingly easy! So, are you question? Just exercise just what we give below as capably as review **Pdf Free Action In Spark** what you as soon as to read!

---

## KEY=ACTION - ARI THOMAS

---

---

### SPARK IN ACTION

---

**Simon and Schuster** Summary Spark in Action teaches you the theory and skills you need to effectively handle batch and streaming data using Spark. Fully updated for Spark 2.0. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Big data systems distribute datasets across clusters of machines, making it a challenge to efficiently query, stream, and interpret them. Spark can help. It is a processing system designed specifically for distributed data. It provides easy-to-use interfaces, along with the performance you need for production-quality analytics and machine learning. Spark 2 also adds improved programming APIs, better performance, and countless other upgrades. About the Book Spark in Action teaches you the theory and skills you need to effectively handle batch and streaming data using Spark. You'll get comfortable with the Spark CLI as you work through a few introductory examples. Then, you'll start programming Spark using its core APIs. Along the way, you'll work with structured data using Spark SQL, process near-real-time streaming data, apply machine learning algorithms, and munge graph data using Spark GraphX. For a zero-effort startup, you can download the preconfigured virtual machine ready for you to try the book's code. What's Inside Updated for Spark 2.0 Real-life case studies Spark DevOps with Docker Examples in Scala, and online in Java and Python About the Reader Written for experienced programmers with some background in big data or machine learning. About the Authors Petar Zečević and Marko Bonaći are seasoned developers heavily involved in the Spark community. Table of Contents PART 1 - FIRST STEPS Introduction to Apache Spark Spark fundamentals Writing Spark applications The Spark API in depth PART 2 - MEET THE SPARK FAMILY Sparkling queries with Spark SQL Ingesting data with Spark Streaming Getting smart with MLlib ML: classification and clustering Connecting the dots with GraphX PART 3 - SPARK OPS Running Spark Running on a Spark standalone cluster Running on YARN and Mesos PART 4 - BRINGING IT TOGETHER Case study: real-time dashboard Deep learning on Spark with H2O

---

### SPARK IN ACTION

---

---

### COVERS APACHE SPARK 3 WITH EXAMPLES IN JAVA, PYTHON, AND SCALA

---

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

[Building a simple app for deployment](#) [6 Deploying your simple app](#) [PART 2 - INGESTION](#) [7 Ingestion from files](#) [8 Ingestion from databases](#) [9 Advanced ingestion: finding data sources and building your own](#) [10 Ingestion through structured streaming](#) [PART 3 - TRANSFORMING YOUR DATA](#) [11 Working with SQL](#) [12 Transforming your data](#) [13 Transforming entire documents](#) [14 Extending transformations with user-defined functions](#) [15 Aggregating your data](#) [PART 4 - GOING FURTHER](#) [16 Cache and checkpoint: Enhancing Spark's performances](#) [17 Exporting data and building full data pipelines](#) [18 Exploring deployment](#)

---

## LEARNING SPARK

---

**O'Reilly Media** Data is bigger, arrives faster, and comes in a variety of formats—and it all needs to be processed at scale for analytics or machine learning. But how can you process such varied workloads efficiently? Enter Apache Spark. Updated to include Spark 3.0, this second edition shows data engineers and data scientists why structure and unification in Spark matters. Specifically, this book explains how to perform simple and complex data analytics and employ machine learning algorithms. Through step-by-step walk-throughs, code snippets, and notebooks, you'll be able to: [Learn Python, SQL, Scala, or Java high-level Structured APIs](#) [Understand Spark operations and SQL Engine](#) [Inspect, tune, and debug Spark operations with Spark configurations and Spark UI](#) [Connect to data sources: JSON, Parquet, CSV, Avro, ORC, Hive, S3, or Kafka](#) [Perform analytics on batch and streaming data using Structured Streaming](#) [Build reliable data pipelines with open source Delta Lake and Spark](#) [Develop machine learning pipelines with MLlib and productionize models using MLflow](#)

---

## SPARK: THE DEFINITIVE GUIDE

---



---

### BIG DATA PROCESSING MADE SIMPLE

---

**"O'Reilly Media, Inc."** Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You'll explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark [Learn about DataFrames, SQL, and Datasets—Spark's core APIs—through worked examples](#) [Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames](#) [Understand how Spark runs on a cluster](#) [Debug, monitor, and tune Spark clusters and applications](#) [Learn the power of Structured Streaming, Spark's stream-processing engine](#) [Learn how you can apply MLlib to a variety of problems, including classification or recommendation](#)

---

## SPARK IN ACTION, SECOND EDITION

---



---

### COVERS APACHE SPARK 3 WITH EXAMPLES IN JAVA, PYTHON, AND SCALA

---

**Manning Publications** Summary The Spark distributed data processing platform provides an easy-to-implement tool for ingesting, streaming, and processing data from any source. In Spark in Action, Second Edition, you'll learn to take advantage of Spark's core features and incredible processing speed, with applications including real-time computation, delayed evaluation, and machine learning. Spark skills are a hot commodity in enterprises worldwide, and with Spark's powerful and flexible Java APIs, you can reap all the benefits without first learning Scala or Hadoop. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology [Analyzing enterprise data starts by reading, filtering, and merging files and streams from many sources. The Spark data processing engine handles this varied volume like a champ, delivering speeds 100 times faster than Hadoop systems. Thanks to SQL support, an intuitive interface, and a straightforward multilanguage API, you can use Spark without learning a complex new ecosystem.](#) About the book [Spark in Action, Second Edition, teaches you to create end-to-end analytics applications. In this entirely new book, you'll learn from interesting Java-based examples, including a complete data pipeline for processing NASA satellite data. And you'll discover Java, Python, and Scala code samples hosted on GitHub that you can explore and adapt, plus appendixes that give you a cheat sheet for installing tools and understanding Spark-specific terms. What's inside](#) [Writing Spark applications in Java](#) [Spark application architecture](#) [Ingestion through files, databases, streaming, and Elasticsearch](#) [Querying distributed datasets with Spark SQL](#) About the reader [This book does not assume previous experience with Spark, Scala, or Hadoop.](#) About the author [Jean-Georges Perrin is an experienced data and software architect. He is France's first IBM Champion and has been honored for 12 consecutive years.](#) Table of Contents [PART 1 - THE THEORY CRIPPLED BY AWESOME EXAMPLES](#) [1 So, what is Spark, anyway?](#) [2 Architecture and flow](#) [3 The majestic role of the dataframe](#) [4 Fundamentally lazy](#) [5 Building a simple app for deployment](#) [6 Deploying your simple app](#) [PART 2 - INGESTION](#) [7 Ingestion from files](#) [8 Ingestion from databases](#) [9 Advanced ingestion: finding data sources and building your own](#) [10 Ingestion through structured streaming](#) [PART 3 - TRANSFORMING YOUR DATA](#) [11 Working with SQL](#) [12 Transforming your data](#) [13 Transforming entire documents](#) [14 Extending transformations with user-defined functions](#) [15 Aggregating your data](#) [PART 4 - GOING FURTHER](#) [16 Cache and checkpoint: Enhancing Spark's performances](#) [17 Exporting data and building full data pipelines](#) [18 Exploring](#)

deployment

---

## SPARK GRAPHX IN ACTION

---

**Simon and Schuster** Summary Spark GraphX in Action starts out with an overview of Apache Spark and the GraphX graph processing API. This example-based tutorial then teaches you how to configure GraphX and how to use it interactively. Along the way, you'll collect practical techniques for enhancing applications and applying machine learning algorithms to graph data. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology GraphX is a powerful graph processing API for the Apache Spark analytics engine that lets you draw insights from large datasets. GraphX gives you unprecedented speed and capacity for running massively parallel and machine learning algorithms. About the Book Spark GraphX in Action begins with the big picture of what graphs can be used for. This example-based tutorial teaches you how to use GraphX interactively. You'll start with a crystal-clear introduction to building big data graphs from regular data, and then explore the problems and possibilities of implementing graph algorithms and architecting graph processing pipelines. Along the way, you'll collect practical techniques for enhancing applications and applying machine learning algorithms to graph data. What's Inside Understanding graph technology Using the GraphX API Developing algorithms for big graphs Machine learning with graphs Graph visualization About the Reader Readers should be comfortable writing code. Experience with Apache Spark and Scala is not required. About the Authors Michael Malak has worked on Spark applications for Fortune 500 companies since early 2013. Robin East has worked as a consultant to large organizations for over 15 years and is a data scientist at Worldpay. Table of Contents PART 1 SPARK AND GRAPHS Two important technologies: Spark and graphs GraphX quick start Some fundamentals PART 2 CONNECTING VERTICES GraphX Basics Built-in algorithms Other useful graph algorithms Machine learning PART 3 OVER THE ARC The missing algorithms Performance and monitoring Other languages and tools

---

## GRAPH ALGORITHMS

---

---

### PRACTICAL EXAMPLES IN APACHE SPARK AND NEO4J

---

**"O'Reilly Media, Inc."** Discover how graph algorithms can help you leverage the relationships within your data to develop more intelligent solutions and enhance your machine learning models. You'll learn how graph analytics are uniquely suited to unfold complex structures and reveal difficult-to-find patterns lurking in your data. Whether you are trying to build dynamic network models or forecast real-world behavior, this book illustrates how graph algorithms deliver value—from finding vulnerabilities and bottlenecks to detecting communities and improving machine learning predictions. This practical book walks you through hands-on examples of how to use graph algorithms in Apache Spark and Neo4j—two of the most common choices for graph analytics. Also included: sample code and tips for over 20 practical graph algorithms that cover optimal pathfinding, importance through centrality, and community detection. Learn how graph analytics vary from conventional statistical analysis Understand how classic graph algorithms work, and how they are applied Get guidance on which algorithms to use for different types of questions Explore algorithm examples with working code and sample datasets from Spark and Neo4j See how connected feature extraction can increase machine learning accuracy and precision Walk through creating an ML workflow for link prediction combining Neo4j and Spark

---

## APACHE SPARK FOR THE ENTERPRISE: SETTING THE BUSINESS FREE

---

**IBM Redbooks** Analytics is increasingly an integral part of day-to-day operations at today's leading businesses, and transformation is also occurring through huge growth in mobile and digital channels. Enterprise organizations are attempting to leverage analytics in new ways and transition existing analytics capabilities to respond with more flexibility while making the most efficient use of highly valuable data science skills. The recent growth and adoption of Apache Spark as an analytics framework and platform is very timely and helps meet these challenging demands. The Apache Spark environment on IBM z/OS® and Linux on IBM z Systems™ platforms allows this analytics framework to run on the same enterprise platform as the originating sources of data and transactions that feed it. If most of the data that will be used for Apache Spark analytics, or the most sensitive or quickly changing data is originating on z/OS, then an Apache Spark z/OS based environment will be the optimal choice for performance, security, and governance. This IBM® Redpaper™ publication explores the enterprise analytics market, use of Apache Spark on IBM z Systems™ platforms, integration between Apache Spark and other enterprise data sources, and case studies and examples of what can be achieved with Apache Spark in enterprise environments. It is of interest to data scientists, data engineers, enterprise architects, or anybody looking to better understand how to combine an analytics framework and platform on enterprise systems.

---

## HIGH PERFORMANCE SPARK

---

---

### BEST PRACTICES FOR SCALING AND OPTIMIZING APACHE SPARK

---

**"O'Reilly Media, Inc."** Apache Spark is amazing when everything clicks. But if you haven't seen the performance improvements you expected, or still don't feel confident enough to use Spark in

production, this practical book is for you. Authors Holden Karau and Rachel Warren demonstrate performance optimizations to help your Spark queries run faster and handle larger data sizes, while using fewer resources. Ideal for software engineers, data engineers, developers, and system administrators working with large-scale data applications, this book describes techniques that can reduce data infrastructure costs and developer hours. Not only will you gain a more comprehensive understanding of Spark, you'll also learn how to make it sing. With this book, you'll explore: How Spark SQL's new interfaces improve performance over SQL's RDD data structure The choice between data joins in Core Spark and Spark SQL Techniques for getting the most out of standard RDD transformations How to work around performance issues in Spark's key/value pair paradigm Writing high-performance Spark code without Scala or the JVM How to test for functionality and performance when applying suggested improvements Using Spark MLlib and Spark ML machine learning libraries Spark's Streaming components and external community packages

---

## LEARNING SPARK

---

### LIGHTNING-FAST BIG DATA ANALYSIS

---

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

---

## BEGINNING APACHE SPARK 2

---

### WITH RESILIENT DISTRIBUTED DATASETS, SPARK SQL, STRUCTURED STREAMING AND SPARK MACHINE LEARNING LIBRARY

---

**Apress** 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.

---

## SPARK GRAPHX IN ACTION

---

**Manning Publications** While graphs are often the most natural way to represent the connections among data, the complexity of large graphs makes them conceptually difficult and computationally expensive to explore, query, and analyze. GraphX, a powerful graph processing API for the Apache Spark analytics engine, makes it possible to efficiently explore and interpret large-scale graph data at near-realtime speeds. GraphX works with Spark's in-memory distributed framework to offer unprecedented speed and capacity for analyzing social media data, performing complex textual analysis, handling important machine learning algorithms, and much more. Spark GraphX in Action starts out with an overview of Apache Spark and the GraphX graph processing API. This example-based tutorial explains how to configure GraphX and use GraphX interactively. It offers a crystal-clear introduction to graph elements, which are needed to build big data graphs. Then, it explores the problems and possibilities of graph algorithm implementations. Along the way, it details practical techniques for enhancing applications and applying machine learning algorithms to graph data. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

---

## THE PRIME OF MISS JEAN BRODIE

---

---

## A NOVEL

---

**Open Road Media** “A perfect book”—and basis for the Maggie Smith film—about a teacher who makes a lasting impression on her female students in the years before World War II (Chicago Tribune). “Give me a girl at an impressionable age, and she is mine for life!” So asserts Jean Brodie, a magnetic, dubious, and sometimes comic teacher at the conservative Marcia Blaine School for Girls in Edinburgh. Brodie selects six favorite pupils to mold—and she doesn’t stop with just their intellectual lives. She has a plan for them all, including how they will live, whom they will love, and what sacrifices they will make to uphold her ideals. When the girls reach adulthood and begin to find their own destinies, Jean Brodie’s indelible imprint is a gift to some, and a curse to others. The Prime of Miss Jean Brodie is Spark’s masterpiece, a novel that offers one of twentieth-century English literature’s most iconic and complex characters—a woman at once admirable and sinister, benevolent and conniving. This ebook features an illustrated biography of Muriel Spark including rare photos and never-before-seen documents from the author’s archive at the National Library of Scotland.

---

## SPARK CITY

---

### BOOK ONE OF THE SPARK CITY CYCLE (LARGE PRINT)

---

Erroh has a plan. A simple plan. It'll never work. Despite his family's warrior pedigree, he'd rather gamble and drink while living from one tavern to the next. But when his wanderings bring him upon a gruesome slaughter of innocents he is torn from carefree ways. Spark City is on the horizon and with it the mysterious trials of The Cull. After a life spent rejecting his birth right, the time has come to pick up his sword and accept his destiny. With an army marching forward, and unlikely companions buy his side, does Erroh have what it takes to stop the coming war? Spark City is the gripping first novel in Robert J Power's best selling epic fantasy series. If you like spectacular combat, unlikely heroes, and slow burn romance, you'll love this thrilling tale. Now availalbe in Large Print Buy Spark City to embrace your calling today!

---

## PRACTICAL APACHE SPARK

---

### USING THE SCALA API

---

**Apress** Work with Apache Spark using Scala to deploy and set up single-node, multi-node, and high-availability clusters. This book discusses various components of Spark such as Spark Core, DataFrames, Datasets and SQL, Spark Streaming, Spark MLlib, and R on Spark with the help of practical code snippets for each topic. Practical Apache Spark also covers the integration of Apache Spark with Kafka with examples. You’ll follow a learn-to-do-by-yourself approach to learning - learn the concepts, practice the code snippets in Scala, and complete the assignments given to get an overall exposure. On completion, you’ll have knowledge of the functional programming aspects of Scala, and hands-on expertise in various Spark components. You’ll also become familiar with machine learning algorithms with real-time usage. What You Will LearnDiscover the functional programming features of Scala Understand the complete architecture of Spark and its componentsIntegrate Apache Spark with Hive and Kafka Use Spark SQL, DataFrames, and Datasets to process data using traditional SQL queries Work with different machine learning concepts and libraries using Spark’s MLlib packages Who This Book Is For Developers and professionals who deal with batch and stream data processing.

---

## MASTERING SPARK WITH R

---

### THE COMPLETE GUIDE TO LARGE-SCALE ANALYSIS AND MODELING

---

**"O'Reilly Media, Inc."** If you’re like most R users, you have deep knowledge and love for statistics. But as your organization continues to collect huge amounts of data, adding tools such as Apache Spark makes a lot of sense. With this practical book, data scientists and professionals working with large-scale data applications will learn how to use Spark from R to tackle big data and big compute problems. Authors Javier Luraschi, Kevin Kuo, and Edgar Ruiz show you how to use R with Spark to solve different data analysis problems. This book covers relevant data science topics, cluster computing, and issues that should interest even the most advanced users. Analyze, explore, transform, and visualize data in Apache Spark with R Create statistical models to extract information and predict outcomes; automate the process in production-ready workflows Perform analysis and modeling across many machines using distributed computing techniques Use large-scale data from multiple sources and different formats with ease from within Spark Learn about alternative modeling frameworks for graph processing, geospatial analysis, and genomics at scale Dive into advanced topics including custom transformations, real-time data processing, and creating custom Spark extensions

---

---

## SOCIAL MEDIA MEASUREMENT AND MANAGEMENT

---

### ENTREPRENEURIAL DIGITAL ANALYTICS

---

**Routledge** This new textbook applies a critical and practical lens to the world of social media analytics. Author Jeremy Harris Lipschultz explores the foundations of digital data, strategic tools, and best practices in an accessible volume for students and practitioners of social media communication. The book expands upon entrepreneurship, marketing, and technological principles, demonstrating how raising awareness, sparking engagement, and producing business outcomes all require emphasis on customers, employees, and other stakeholders within paid, earned, social, and owned media. It also looks to the future, examining how the movement toward artificial intelligence and machine learning raises new legal and ethical issues in effective management of social media data. Additionally, the book offers a solid grounding in the principles of social media measurement itself, teaching the strategies and techniques that enable effective analysis. A perfect primer for this developing industry, *Social Media Measurement: Entrepreneurial Digital Analytics* is ideal for students, scholars, and practitioners of digital media seeking to hone their skills and expand their bank of tools and resources. It features theoretical and practical advice, a comprehensive glossary of key terms, and case studies from key industry thought leaders.

---

### SPARK

---

### THE REVOLUTIONARY NEW SCIENCE OF EXERCISE AND THE BRAIN

---

**Little, Brown Spark** An investigation into the effects of exercise on the brain evaluates how aerobic exercise positively influences the progression of such conditions as Alzheimer's disease, ADD, and depression, in a report that shares theory-supporting case studies and the results of a progressive school fitness program. 30,000 first printing.

---

### EVERYDAY USE

---

**Rutgers University Press** Presents the text of Alice Walker's story "Everyday Use"; contains background essays that provide insight into the story; and features a selection of critical response. Includes a chronology and an interview with the author.

---

### DATA ALGORITHMS WITH SPARK

---

**"O'Reilly Media, Inc."** Apache Spark's speed, ease of use, sophisticated analytics, and multilanguage support makes practical knowledge of this cluster-computing framework a required skill for data engineers and data scientists. With this hands-on guide, anyone looking for an introduction to Spark will learn practical algorithms and examples using PySpark. In each chapter, author Mahmoud Parsian shows you how to solve a data problem with a set of Spark transformations and algorithms. You'll learn how to tackle problems involving ETL, design patterns, machine learning algorithms, data partitioning, and genomics analysis. Each detailed recipe includes PySpark algorithms using the PySpark driver and shell script. With this book, you will: Learn how to select Spark transformations for optimized solutions Explore powerful transformations and reductions including `reduceByKey()`, `combineByKey()`, and `mapPartitions()` Understand data partitioning for optimized queries Build and apply a model using PySpark design patterns Apply motif-finding algorithms to graph data Analyze graph data by using the GraphFrames API Apply PySpark algorithms to clinical and genomics data Learn how to use and apply feature engineering in ML algorithms Understand and use practical and pragmatic data design patterns

---

### GALVANI'S SPARK

---

### THE STORY OF THE NERVE IMPULSE

---

**Oxford University Press** Galvani's Spark describes the gradual understanding of the nerve impulse--the basis of all thoughts, sensations and actions. Proceeding from a chance observation on a frog leg, to studies on squid giant axons and bacteria, this book concludes with the increasing realization that ion channels are responsible for a variety of clinical disorders. Written by a neurophysiologist, the book provides portraits of some of the greatest names in medical science.

---

### A WRINKLE IN TIME

---

**GENERAL PRESS** *A Wrinkle in Time* is the winner of the 1963 Newbery Medal. It was a dark and stormy night—Meg Murry, her small brother Charles Wallace, and her mother had come down to the kitchen for a midnight snack when they were upset by the arrival of a most disturbing stranger. "Wild nights are my glory," the unearthly stranger told them. "I just got caught in a downdraft and blown off

course. Let me sit down for a moment, and then I'll be on my way. Speaking of ways, by the way, there is such a thing as a tesseract." A tesseract (in case the reader doesn't know) is a wrinkle in time. To tell more would rob the reader of the enjoyment of Miss L'Engle's unusual book. A Wrinkle in Time, winner of the Newbery Medal in 1963, is the story of the adventures in space and time of Meg, Charles Wallace, and Calvin O'Keefe (athlete, student, and one of the most popular boys in high school). They are in search of Meg's father, a scientist who disappeared while engaged in secret work for the government on the tesseract problem.

---

## DATA ANALYTICS WITH HADOOP

---

---

### AN INTRODUCTION FOR DATA SCIENTISTS

---

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

---

### PRO SPARK STREAMING

---

---

#### THE ZEN OF REAL-TIME ANALYTICS USING APACHE SPARK

---

**Apress** Learn the right cutting-edge skills and knowledge to leverage Spark Streaming to implement a wide array of real-time, streaming applications. This book walks you through end-to-end real-time application development using real-world applications, data, and code. Taking an application-first approach, each chapter introduces use cases from a specific industry and uses publicly available datasets from that domain to unravel the intricacies of production-grade design and implementation. The domains covered in Pro Spark Streaming include social media, the sharing economy, finance, online advertising, telecommunication, and IoT. In the last few years, Spark has become synonymous with big data processing. DStreams enhance the underlying Spark processing engine to support streaming analysis with a novel micro-batch processing model. Pro Spark Streaming by Zubair Nabi will enable you to become a specialist of latency sensitive applications by leveraging the key features of DStreams, micro-batch processing, and functional programming. To this end, the book includes ready-to-deploy examples and actual code. Pro Spark Streaming will act as the bible of Spark Streaming. What You'll Learn Discover Spark Streaming application development and best practices Work with the low-level details of discretized streams Optimize production-grade deployments of Spark Streaming via configuration recipes and instrumentation using Graphite, collectd, and Nagios Ingest data from disparate sources including MQTT, Flume, Kafka, Twitter, and a custom HTTP receiver Integrate and couple with HBase, Cassandra, and Redis Take advantage of design patterns for side-effects and maintaining state across the Spark Streaming micro-batch model Implement real-time and scalable ETL using data frames, SparkSQL, Hive, and SparkR Use streaming machine learning, predictive analytics, and recommendations Mesh batch processing with stream processing via the Lambda architecture Who This Book Is For Data scientists, big data experts, BI analysts, and data architects.

---

### A DOLL'S HOUSE

---

**BoD - Books on Demand** Reproduction of the original: A Doll's House by Henrik Ibsen

---

### DATA ALGORITHMS

---

---

#### RECIPES FOR SCALING UP WITH HADOOP AND SPARK

---

**"O'Reilly Media, Inc."** If you are ready to dive into the MapReduce framework for processing large datasets, this practical book takes you step by step through the algorithms and tools you need to build distributed MapReduce applications with Apache Hadoop or Apache Spark. Each chapter provides a recipe for solving a massive computational problem, such as building a recommendation system. You'll learn how to implement the appropriate MapReduce solution with code that you can use in your projects. Dr. Mahmoud Parsian covers basic design patterns, optimization techniques, and data mining and

machine learning solutions for problems in bioinformatics, genomics, statistics, and social network analysis. This book also includes an overview of MapReduce, Hadoop, and Spark. Topics include: Market basket analysis for a large set of transactions Data mining algorithms (K-means, KNN, and Naive Bayes) Using huge genomic data to sequence DNA and RNA Naive Bayes theorem and Markov chains for data and market prediction Recommendation algorithms and pairwise document similarity Linear regression, Cox regression, and Pearson correlation Allelic frequency and mining DNA Social network analysis (recommendation systems, counting triangles, sentiment analysis)

---

## THE ILIAD

---

**Aegitas** The Iliad is an ancient Greek epic poem in dactylic hexameter, traditionally attributed to Homer. Set during the Trojan War, the ten-year siege of the city of Troy (Ilium) by a coalition of Greek states, it tells of the battles and events during the weeks of a quarrel between King Agamemnon and the warrior Achilles.

---

## BEING THE CHANGE

---



---

### LIVE WELL AND SPARK A CLIMATE REVOLUTION

---

**New Society Publishers** Life on 1/10th the fossil fuels turns out to be awesome. We all want to be happy. Yet as we consume ever more in a frantic bid for happiness, global warming worsens. Alarmed by drastic changes now occurring in the Earth's climate systems, the author, a climate scientist and suburban father of two, embarked on a journey to change his life and the world. He began by bicycling, growing food, meditating, and making other simple, fulfilling changes. Ultimately, he slashed his climate impact to under a tenth of the US average and became happier in the process. Being the Change explores the connections between our individual daily actions and our collective predicament. It merges science, spirituality, and practical action to develop a satisfying and appropriate response to global warming. Part one exposes our interconnected predicament: overpopulation, global warming, industrial agriculture, growth-addicted economics, a sold-out political system, and a mindset of separation from nature. It also includes a readable but authoritative overview of climate science. Part two offers a response at once obvious and unprecedented: mindfully opting out of this broken system and aligning our daily lives with the biosphere. The core message is deeply optimistic: living without fossil fuels is not only possible, it can be better. Peter Kalmus is an atmospheric scientist at Caltech / Jet Propulsion Laboratory with a Ph.D. in physics from Columbia University. He lives in suburban Altadena, California with his wife and two children on 1/10th the fossil fuels of the average American. Peter speaks purely on his own behalf, not on behalf of NASA or Jet Propulsion Laboratory.

---

## NICOMACHEAN ETHICS

---

**Courier Corporation** Enduringly profound treatise, whose lasting effect on Western philosophy continues to resonate. Aristotle identifies the goal of life as happiness and discusses its attainment through the contemplation of philosophic truth.

---

## THINGS FALL APART

---

**Penguin UK** Okonkwo is the greatest warrior alive, famous throughout West Africa. But when he accidentally kills a clansman, things begin to fall apart. Then Okonkwo returns from exile to find missionaries and colonial governors have arrived in the village. With his world thrown radically off-balance he can only hurtle towards tragedy. Chinua Achebe's stark novel reshaped both African and world literature. This arresting parable of a proud but powerless man witnessing the ruin of his people begins Achebe's landmark trilogy of works chronicling the fate of one African community, continued in Arrow of God and No Longer at Ease.

---

## HISTORY OF PLYMOUTH PLANTATION, 1620-1647

---



---

## MODERN WARFARE

---



---

## A FRENCH VIEW OF COUNTERINSURGENCY

---

**DIANE Publishing**

---

---

## THE BOOK THIEF

---

**Knopf Books for Young Readers** #1 NEW YORK TIMES BESTSELLER • ONE OF TIME MAGAZINE'S 100 BEST YA BOOKS OF ALL TIME The extraordinary, beloved novel about the ability of books to feed the soul even in the darkest of times. When Death has a story to tell, you listen. It is 1939. Nazi Germany. The country is holding its breath. Death has never been busier, and will become busier still. Liesel Meminger is a foster girl living outside of Munich, who scratches out a meager existence for herself by stealing when she encounters something she can't resist—books. With the help of her accordion-playing foster father, she learns to read and shares her stolen books with her neighbors during bombing raids as well as with the Jewish man hidden in her basement. In superbly crafted writing that burns with intensity, award-winning author Markus Zusak, author of *I Am the Messenger*, has given us one of the most enduring stories of our time. "The kind of book that can be life-changing." —The New York Times "Deserves a place on the same shelf with *The Diary of a Young Girl* by Anne Frank." —USA Today DON'T MISS BRIDGE OF CLAY, MARKUS ZUSAK'S FIRST NOVEL SINCE THE BOOK THIEF.

---

---

## METHODS OF SOIL ANALYSIS, PART 3

---

---

### CHEMICAL METHODS

---

**John Wiley & Sons** A thorough presentation of analytical methods for characterizing soil chemical properties and processes, *Methods, Part 3* includes chapters on Fourier transform infrared, Raman, electron spin resonance, x-ray photoelectron, and x-ray absorption fine structure spectroscopies, and more.

---

---

## WAR OF THE SPARK: RAVNICA (MAGIC: THE GATHERING)

---

**Del Rey** NEW YORK TIMES BESTSELLER • Experience the first official adventure in *Magic: The Gathering's* multiverse in nearly a decade as the ultimate battle begins on Ravnica. Teyo Verada wants nothing more than to be a shieldmage, wielding arcane energies to protect his people from his world's vicious diamondstorms. When he's buried alive in the aftermath of his first real tempest, the young mage's life is about to end before it can truly begin—until it doesn't. In a flash, a power he didn't know he had whisks him away from his home, to a world of stone, glass, and wonder: Ravnica. Teyo is a Planeswalker, one of many to be called to the world-spanning city—all lured by Nicol Bolas, the Elder Dragon. Bolas lays siege to the city of Ravnica, hungry for the ultimate prize: godhood itself. His unparalleled magic and unstoppable army appear poised to bring the city to utter ruin. Among those who stand in the way of Bolas's terrifying machinations are the Gatewatch, Planeswalkers sworn to defeat evil, no matter where it's found. But as they work to unite the other mages and mount a defense of the city and its people, the terrifying truth of Bolas's plan becomes clear. The Elder Dragon has prepared a trap to ensnare the most powerful mages from across the Multiverse—and it's too late to escape. As forces great and small converge on the city and the battle rages, the stakes could not be higher. If the Gatewatch falters and the Planeswalkers fail, the curtain will fall on the age of heroes—and rise on the infinite reign of Nicol Bolas.

---

---

## SILAS MARNER ILLUSTRATED

---

*Silas Marner* is the third novel by George Eliot, published in 1861. An outwardly simple tale of a linen weaver, it is notable for its strong realism and its sophisticated treatment of a variety of issues ranging from religion to industrialisation to community.

---

---

## THE 5 SECOND RULE

---

---

### TRANSFORM YOUR LIFE, WORK, AND CONFIDENCE WITH EVERYDAY COURAGE

---

**Savio Republic** Throughout your life, you've had parents, coaches, teachers, friends, and mentors who have pushed you to be better than your excuses and bigger than your fears. What if the secret to having the confidence and courage to enrich your life and work is simply knowing how to push yourself? Using the science habits, riveting stories and surprising facts from some of the most famous moments in history, art and business, Mel Robbins will explain the power of a "push moment." Then, she'll give you one simple tool you can use to become your greatest self. It takes just five seconds to use this tool, and every time you do, you'll be in great company. More than 8 million people have watched Mel's TEDx Talk, and executives inside of the world's largest brands are using the tool to increase productivity, collaboration, and engagement. In *The 5 Second Rule*, you'll discover it takes just five seconds to: Become confident Break the habit of procrastination and self-doubt Beat fear and uncertainty Stop worrying and feel happier Share your ideas with courage *The 5 Second Rule* is a simple, one-size-fits-all solution for the one problem we all face—we hold ourselves back. The secret isn't knowing what to do—it's knowing how to make yourself do it. p.p1 {margin: 0.0px 0.0px 0.0px 0.0px; font: 12.0px Arial}

---

---

**SOPHIE'S WORLD**

---

**A NOVEL ABOUT THE HISTORY OF PHILOSOPHY**

---

**Farrar, Straus and Giroux** One day Sophie comes home from school to find two questions in her mail: "Who are you?" and "Where does the world come from?" Before she knows it she is enrolled in a correspondence course with a mysterious philosopher. Thus begins Jostein Gaarder's unique novel, which is not only a mystery, but also a complete and entertaining history of philosophy.

---

**TEARS OF A TIGER**

---

**Simon and Schuster** Three boys struggle to come to terms with the death of a friend in a drunk-driving auto accident in which all four were involved, in a story told through newspaper stories, diary entries, school announcements, telephone conversations, and classroom assignments.

---

**THE BOOK OF UNKNOWN AMERICANS**

---

**A NOVEL**

---

**Vintage** "Illuminate[s] the lives behind the current debates about Latino immigration." —The New York Times Book Review When fifteen-year-old Maribel Rivera sustains a terrible injury, the Riveras leave behind a comfortable life in Mexico and risk everything to come to the United States so that Maribel can have the care she needs. Once they arrive, it's not long before Maribel attracts the attention of Mayor Toro, the son of one of their new neighbors, who sees a kindred spirit in this beautiful, damaged outsider. Their love story sets in motion events that will have profound repercussions for everyone involved. Here Henríquez seamlessly interweaves the story of these star-crossed lovers, and of the Rivera and Toro families, with the testimonials of men and women who have come to the United States from all over Latin America. The Book of Unknown Americans is a stunning novel of hopes and dreams, guilt and love—a book that offers a resonant new definition of what it means to be American. Named a New York Times and Washington Post Notable Book, an NPR Great Read, The Daily Beast's Novel of the Year, and a Mother Jones, Oprah.com, School Library Journal, and BookPage Best Book of the Year