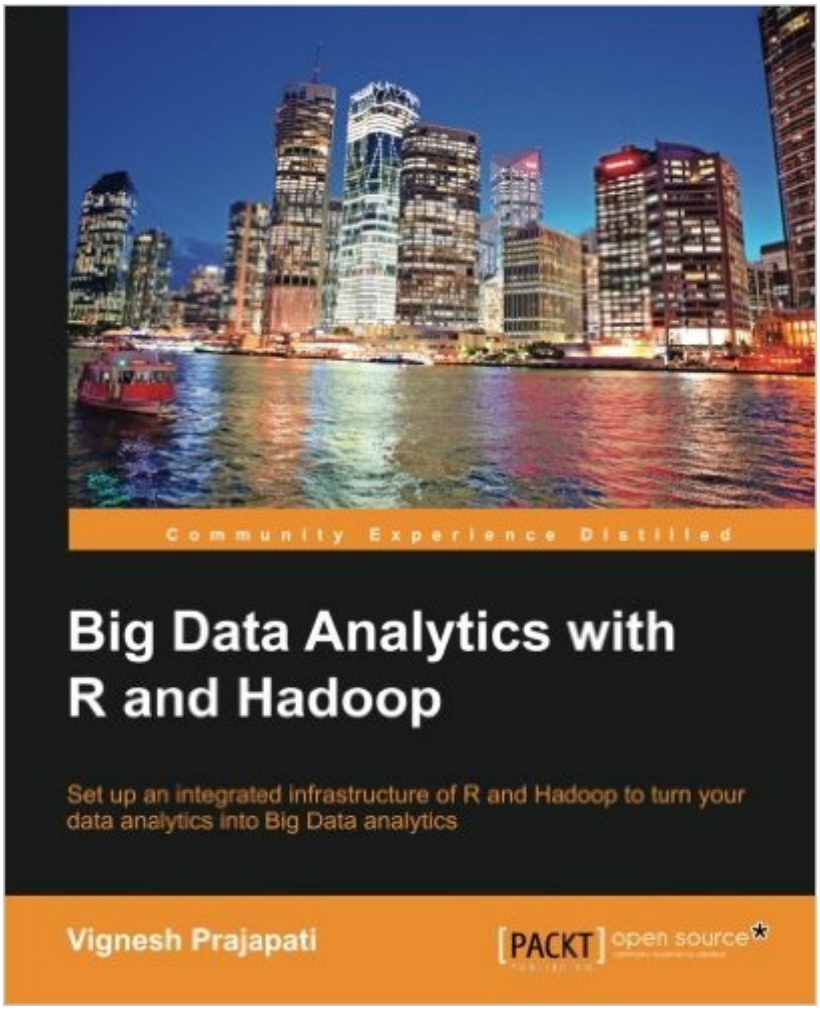
The book was found

Big Data Analytics With R And Hadoop



PDF
Adobe
DOWNLOAD EBOOK

Synopsis

If you’re an R developer looking to harness the power of big data analytics with Hadoop, then this book tells you everything you need to integrate the two. You’ll end up capable of building a data analytics engine with huge potential. Overview Write Hadoop MapReduce within R Learn data analytics with R and the Hadoop platform Handle HDFS data within R Understand Hadoop streaming with R Encode and enrich datasets into R In Detail Big data analytics is the process of examining large amounts of data of a variety of types to uncover hidden patterns, unknown correlations, and other useful information. Such information can provide competitive advantages over rival organizations and result in business benefits, such as more effective marketing and increased revenue. New methods of working with big data, such as Hadoop and MapReduce, offer alternatives to traditional data warehousing. Big Data Analytics with R and Hadoop is focused on the techniques of integrating R and Hadoop by various tools such as RHIPE and RHadoop. A powerful data analytics engine can be built, which can process analytics algorithms over a large scale dataset in a scalable manner. This can be implemented through data analytics operations of R, MapReduce, and HDFS of Hadoop. You will start with the installation and configuration of R and Hadoop. Next, you will discover information on various practical data analytics examples with R and Hadoop. Finally, you will learn how to import/export from various data sources to R. Big Data Analytics with R and Hadoop will also give you an easy understanding of the R and Hadoop connectors RHIPE, RHadoop, and Hadoop streaming. What you will learn from this book Integrate R and Hadoop via RHIPE, RHadoop, and Hadoop streaming Develop and run a MapReduce application that runs with R and Hadoop Handle HDFS data from within R using RHIPE and RHadoop Run Hadoop streaming and MapReduce with R Import and export from various data sources to R Approach Big Data Analytics with R and Hadoop is a tutorial style book that focuses on all the powerful big data tasks that can be achieved by integrating R and Hadoop. Who this book is written for This book is ideal for R developers who are looking for a way to perform big data analytics with Hadoop. This book is also aimed at those who know Hadoop and want to build some intelligent applications over Big data with R packages. It would be helpful if readers have basic knowledge of R.

Book Information

Paperback: 238 pages
Publisher: Packt Publishing (November 25, 2013)
Language: English

ISBN-10: 178216328X
ISBN-13: 978-1782163282
Product Dimensions: 7.5 x 0.5 x 9.2 inches
Shipping Weight: 1.2 pounds (View shipping rates and policies)
Average Customer Review: 3.0 out of 5 starsÂ Â See all reviewsÂ (16 customer reviews)
Best Sellers Rank: #633,125 in Books (See Top 100 in Books) #244 inÂ Books > Computers & Technology > Networking & Cloud Computing > Network Administration > Storage & Retrieval #358 inÂ Books > Computers & Technology > Databases & Big Data > Data Mining #460 inÂ Books > Computers & Technology > Databases & Big Data > Data Processing

Customer Reviews

Here I am, a beginner in need of guidance and happy to reward a helpful getting-started book with praise. Fast-forward one day - it’s not gonna happen. Having installed Hadoop and almost installed RHadoop (oh boy, it is going to take a while before this technology becomes as user-friendly as any Microsoft product), I have found the book’s directions regarding installation in-your-face inadequate - and once I started relying on online references (such as Michael Noll’s Hadoop-on-Ubuntu article and resources linked from Revolution Analytics RHadoop Wiki), I really do not feel like going back to this poorly written and edited mess with perhaps 70 smallish pages of useful content. (RHIPE - not interested. Hadoop Streaming - not interested. Interfacing R with MySQL, etc. - does not belong in the book, and not interested). At $10, the book, despite its flaws, could be entertained as a convenience, but at $20 - not to mention $45 for a paper copy - it’s "Forget about it, Packt".PS. A sentence on page 43 catches my attention."The MapReduce framework is notoriously difficult to leverage for transformational logic that is not as simple, for example, real-time streaming, graph processing, and message passing".The "notoriously difficult to leverage" bit sounds a bit too eloquent when compared to the surrounding text. A quick Google search, and hey-ho,"The MapReduce framework is notoriously difficult to leverage for more than simple transformational logic".says a 2012 white paper by ParAccel Inc. It figures.

I never write reviews, but the several hours I wasted trying to decipher this POS impels me to alert others to avoid this time sink. I have never seen in print such poorly crafted prose even on Wikipedia. The contents here we’re obviously copy-pasted from various web sites and redundantly at that, by someone or some group of folks who have little fluency in English nor Hadoop, R, Big Data ...Seriously I could rate a -1

I got a copy of this book and went through it and i am not impressed. The author has just said whats out there and does not seem to have the real experience of actually worked on these. Its more like a pointer work and i did not find any thing informative or deep learning of sort on concepts. Most info on the book can be got just by googling and reading .org websites intro page..Would not recommend to buy this. 2 star for those simple examples. I felt extreme redundant material when i went from chapter to chapter. donot buy.

I don’t usually write angry reviews, and I am really sorry to do so now, but this book has too many issues to ignore them. I have purchased other books from Packt, and I consider it a serious and good quality company. This is why I really can’t understand how they managed to publish something like this. The English is appalling and the content isn’t any better. Explanations are superficial, unclear or chaotic. Mistakes and meaningless sentences pop up all the time. What’s event worse is that, according to the publisher, this book has been reviewed by four (yes four!) different reviewers. It takes less then 30 minutes to understand that this book needs to be rewritten from scratch, so the natural question is: have these these gentlemen read this book? I seriously doubt it. It is really a missed opportunity, since this book would have filled a very important gap. I am really, really disappointed. Maybe, after massive rewriting and editing, it could become at least acceptable, but as for now, don’t do the mistake I have done purchasing it.

Author could had worked more in this domain. At-times, I felt like this book is a compilation of all the top 10 google search - ’ R , Hadoop and RHadoop’. I could conclude saying that everone who googles about Big data / Hadoop and R can write a book.The best picture in the book was a picture depicting what is RHadoopand the answer is :R + Hadoop = RHadoop

To integrate Hadoop with R seems to be an increasingly popular topic; R is one of the most favoured languages for data analysis and with the rise of big data, it became a logical requirement to be able to use R together with Hadoop. The book covers the options that are available to run R jobs on Hadoop clusters. It describes RHadoop and RHIPE packages as well as Hadoop Streaming. The first four chapters provide the basic information on how to install R and Hadoop, how a simple MapReduce job looks like and a couple of standard wordcount examples. The fifth chapter is about using R and Hadoop in 3 different use cases; I had a sort of mixed feeling about them but I liked the way how they tried to illustrate the different key steps from pre-processing and

loading the data via performing analysis to visualizing the output. I also liked the machine learning examples demonstrating both the R-only solutions as well as using R and Hadoop together. The last chapter is a bit of a mix in my opinion; I do not think that SQLite or Excel really belong to this topic, though Hive/RHive and HBase/RHBase certainly deserve attention when we are talking about big data analysis and R. If your aim is to learn how to run R jobs on Hadoop, this book could provide you useful guidance.

Download to continue reading...

Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data) Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business. Leveraging the Power of Data Analytics, Data ... (Hacking Freedom and Data Driven) (Volume 2) Data Analytics with Hadoop: An Introduction for Data Scientists Agile Data Science: Building Data Analytics Applications with Hadoop Big Data Analytics with R and Hadoop Analytics: Data Science, Data Analysis and Predictive Analytics for Business Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data From Big Data to Big Profits: Success with Data and Analytics Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems Practical Hive: A Guide to Hadoop's Data Warehouse System RapidMiner: Data Mining Use Cases and Business Analytics Applications (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series) Healthcare Data Analytics (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series) Even You Can Learn Statistics and Analytics: An Easy to Understand Guide to Statistics and Analytics (3rd Edition) People Analytics: How Social Sensing Technology Will Transform Business and What It Tells Us about the Future of Work (FT Press Analytics) Cognitive Computing and Big Data Analytics Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications) Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking Hadoop Application Architectures

Dmca