



# Big Data Analytics with Spark and Hadoop

*Venkat Ankam*

Download now

[Click here](#) if your download doesn't start automatically

# Big Data Analytics with Spark and Hadoop

*Venkat Ankam*

**Big Data Analytics with Spark and Hadoop** Venkat Ankam

**A handy reference guide for data analysts and data scientists to fetch “Value” out of big data analytics using Spark on Hadoop Clusters**

## About This Book

- Practical tutorial with real-world examples that explores Spark on Hadoop clusters
- This book is based on the latest version of Apache Spark and Hadoop integrated with the most commonly used tools
- Learn about all the Spark stack components including the latest topics such as DataFrames, DataSets, and SparkR

## Who This Book Is For

Though this book is primarily aimed at data analysts and data scientists, it will also help architects, programmers, and practitioners. Knowledge of either Spark or Hadoop would be beneficial. It is assumed that you have basic programming background in Scala, Python, SQL, or R programming with basic Linux experience. Working experience within big data environments is not mandatory.

## What You Will Learn

- Find out about and implement the tools and techniques of big data analytics using Spark on Hadoop clusters
- Understand all the Hadoop and Spark ecosystem components and how Spark replaced MapReduce
- Get to know all the Spark components: Spark Core, Spark SQL, DataFrames, DataSets, Streaming, MLLib, and Graphx
- See batch and real-time data analytics using Spark Core, Spark SQL, and Spark Streaming
- Get to grips with data science and machine learning using MLLib, H2O, Hivemall, Graphx, and SparkR
- Get an introduction to all the new tools (based on Notebooks, Data Flow, and Spark as a Service) and their integrations with Spark and Hadoop

## In Detail

This book explains the fundamentals of Apache Spark and Hadoop, and how they are easily integrated together with the most commonly used tools and techniques. All the Spark components—Spark Core, Spark SQL, DataFrames, Data sets, Streaming, MLLib, Graphx, and Hadoop core components—HDFS, MapReduce, and Yarn are explored in greater depth with implementation examples on Spark and Hadoop clusters.

The big data analytics industry is moving away from MapReduce to Spark. In this book, the advantages of Spark over MapReduce are explained at great depth so you can reap the benefits of in-memory speeds. The DataFrames API, Data Sources API, and new Data sets API are explained so you can build big data

analytical applications.

We'll explore real-time data analytics using Spark Streaming with Apache Kafka and HBase to help you build streaming applications. You'll get to know the machine learning techniques using MLlib and SparkR, and Graph Analytics with the GraphX component of Spark. You will also get the opportunity to start working with web-based notebooks such as Jupyter, Apache Zeppelin, and the data flow tool Apache NiFi to analyze and visualize data.

 [Download Big Data Analytics with Spark and Hadoop ...pdf](#)

 [Read Online Big Data Analytics with Spark and Hadoop ...pdf](#)

## **Download and Read Free Online Big Data Analytics with Spark and Hadoop Venkat Ankam**

---

### **From reader reviews:**

#### **Earl Diehl:**

Nowadays reading books become more and more than want or need but also turn into a life style. This reading habit give you lot of advantages. Advantages you got of course the knowledge the actual information inside the book that will improve your knowledge and information. The info you get based on what kind of guide you read, if you want attract knowledge just go with training books but if you want really feel happy read one using theme for entertaining including comic or novel. Typically the Big Data Analytics with Spark and Hadoop is kind of publication which is giving the reader erratic experience.

#### **Beverly Ingram:**

Are you kind of active person, only have 10 or even 15 minute in your moment to upgrading your mind skill or thinking skill even analytical thinking? Then you are experiencing problem with the book as compared to can satisfy your short time to read it because this all time you only find book that need more time to be examine. Big Data Analytics with Spark and Hadoop can be your answer because it can be read by an individual who have those short extra time problems.

#### **Christopher Bohner:**

A lot of publication has printed but it takes a different approach. You can get it by world wide web on social media. You can choose the very best book for you, science, comedian, novel, or whatever by simply searching from it. It is referred to as of book Big Data Analytics with Spark and Hadoop. You'll be able to your knowledge by it. Without making the printed book, it may add your knowledge and make you actually happier to read. It is most critical that, you must aware about book. It can bring you from one destination to other place.

#### **Joel Padilla:**

What is your hobby? Have you heard which question when you got students? We believe that that issue was given by teacher to the students. Many kinds of hobby, All people has different hobby. Therefore you know that little person just like reading or as reading become their hobby. You should know that reading is very important and also book as to be the issue. Book is important thing to increase you knowledge, except your current teacher or lecturer. You see good news or update concerning something by book. Many kinds of books that can you choose to adopt be your object. One of them is niagra Big Data Analytics with Spark and Hadoop.

**Download and Read Online Big Data Analytics with Spark and Hadoop Venkat Ankam #KX9C7JE8MHF**

## **Read Big Data Analytics with Spark and Hadoop by Venkat Ankam for online ebook**

Big Data Analytics with Spark and Hadoop by Venkat Ankam Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Big Data Analytics with Spark and Hadoop by Venkat Ankam books to read online.

### **Online Big Data Analytics with Spark and Hadoop by Venkat Ankam ebook PDF download**

**Big Data Analytics with Spark and Hadoop by Venkat Ankam Doc**

**Big Data Analytics with Spark and Hadoop by Venkat Ankam Mobipocket**

**Big Data Analytics with Spark and Hadoop by Venkat Ankam EPub**