



# Big Data Analytics

By Venkat Ankam

[Download now](#)

[Read Online](#) 

## Big Data Analytics By Venkat Ankam

### Key Features

- This book is based on the latest 2.0 version of Apache Spark and 2.7 version of Hadoop integrated with most commonly used tools.
- Learn all Spark stack components including latest topics such as DataFrames, DataSets, GraphFrames, Structured Streaming, DataFrame based ML Pipelines and SparkR.
- Integrations with frameworks such as HDFS, YARN and tools such as Jupyter, Zeppelin, NiFi, Mahout, HBase Spark Connector, GraphFrames, H2O and Hivemall.

### Book Description

Big Data Analytics book aims at providing the fundamentals of Apache Spark and Hadoop. All Spark components – Spark Core, Spark SQL, DataFrames, Data sets, Conventional Streaming, Structured Streaming, MLLib, Graphx and Hadoop core components – HDFS, MapReduce and Yarn are explored in greater depth with implementation examples on Spark + Hadoop clusters.

It is moving away from MapReduce to Spark. So, advantages of Spark over MapReduce are explained at great depth to reap benefits of in-memory speeds. DataFrames API, Data Sources API and new Data set API are explained for building Big Data analytical applications. Real-time data analytics using Spark Streaming with Apache Kafka and HBase is covered to help building streaming applications. New Structured streaming concept is explained with an IOT (Internet of Things) use case. Machine learning techniques are covered using MLLib, ML Pipelines and SparkR and Graph Analytics are covered with GraphX and GraphFrames components of Spark.

Readers will also get an opportunity to get started with web based notebooks such as Jupyter, Apache Zeppelin and data flow tool Apache NiFi to analyze and visualize data.

### What you will learn

- Find out and implement the tools and techniques of big data analytics using

Spark on Hadoop clusters with wide variety of tools used with Spark and Hadoop

- Understand all the Hadoop and Spark ecosystem components
- Get to know all the Spark components: Spark Core, Spark SQL, DataFrames, DataSets, Conventional and Structured Streaming, MLLib, ML Pipelines and Graphx
- See batch and real-time data analytics using Spark Core, Spark SQL, and Conventional and Structured Streaming
- Get to grips with data science and machine learning using MLLib, ML Pipelines, H2O, Hivemall, Graphx, SparkR and Hivemall.

### About the Author

**Venkat Ankam** has over 18 years of IT experience and over 5 years in big data technologies, working with customers to design and develop scalable big data applications. Having worked with multiple clients globally, he has tremendous experience in big data analytics using Hadoop and Spark.

He is a Cloudera Certified Hadoop Developer and Administrator and also a Databricks Certified Spark Developer. He is the founder and presenter of a few Hadoop and Spark meetup groups globally and loves to share knowledge with the community.

Venkat has delivered hundreds of trainings, presentations, and white papers in the big data sphere. While this is his first attempt at writing a book, many more books are in the pipeline.

### Table of Contents

1. Big Data Analytics at 10,000 foot view
2. Getting Started with Apache Hadoop and Apache Spark
3. Deep Dive into Apache Spark
4. Big Data Analytics with Spark SQL, DataFrames, and Datasets
5. Real-Time Analytics with Spark Streaming and Structured Streaming
6. Notebooks and Dataflows with Spark and Hadoop
7. Machine Learning with Spark and Hadoop
8. Building Recommendation Systems with Spark and Mahout
9. Graph Analytics with GraphX
10. Interactive Analytics with SparkR



[Download Big Data Analytics ...pdf](#)



[Read Online Big Data Analytics ...pdf](#)

# Big Data Analytics

By Venkat Ankam

## Big Data Analytics By Venkat Ankam

### Key Features

- This book is based on the latest 2.0 version of Apache Spark and 2.7 version of Hadoop integrated with most commonly used tools.
- Learn all Spark stack components including latest topics such as DataFrames, DataSets, GraphFrames, Structured Streaming, DataFrame based ML Pipelines and SparkR.
- Integrations with frameworks such as HDFS, YARN and tools such as Jupyter, Zeppelin, NiFi, Mahout, HBase Spark Connector, GraphFrames, H2O and Hivemall.

### Book Description

Big Data Analytics book aims at providing the fundamentals of Apache Spark and Hadoop. All Spark components – Spark Core, Spark SQL, DataFrames, Data sets, Conventional Streaming, Structured Streaming, MLLib, Graphx and Hadoop core components – HDFS, MapReduce and Yarn are explored in greater depth with implementation examples on Spark + Hadoop clusters.

It is moving away from MapReduce to Spark. So, advantages of Spark over MapReduce are explained at great depth to reap benefits of in-memory speeds. DataFrames API, Data Sources API and new Data set API are explained for building Big Data analytical applications. Real-time data analytics using Spark Streaming with Apache Kafka and HBase is covered to help building streaming applications. New Structured streaming concept is explained with an IOT (Internet of Things) use case. Machine learning techniques are covered using MLLib, ML Pipelines and SparkR and Graph Analytics are covered with GraphX and GraphFrames components of Spark.

Readers will also get an opportunity to get started with web based notebooks such as Jupyter, Apache Zeppelin and data flow tool Apache NiFi to analyze and visualize data.

### What you will learn

- Find out and implement the tools and techniques of big data analytics using Spark on Hadoop clusters with wide variety of tools used with Spark and Hadoop
- Understand all the Hadoop and Spark ecosystem components
- Get to know all the Spark components: Spark Core, Spark SQL, DataFrames, DataSets, Conventional and Structured Streaming, MLLib, ML Pipelines and Graphx
- See batch and real-time data analytics using Spark Core, Spark SQL, and Conventional and Structured Streaming
- Get to grips with data science and machine learning using MLLib, ML Pipelines, H2O, Hivemall, Graphx, SparkR and Hivemall.

### About the Author

**Venkat Ankam** has over 18 years of IT experience and over 5 years in big data technologies, working with customers to design and develop scalable big data applications. Having worked with multiple clients

globally, he has tremendous experience in big data analytics using Hadoop and Spark.

He is a Cloudera Certified Hadoop Developer and Administrator and also a Databricks Certified Spark Developer. He is the founder and presenter of a few Hadoop and Spark meetup groups globally and loves to share knowledge with the community.

Venkat has delivered hundreds of trainings, presentations, and white papers in the big data sphere. While this is his first attempt at writing a book, many more books are in the pipeline.

## Table of Contents

1. Big Data Analytics at 10,000 foot view
2. Getting Started with Apache Hadoop and Apache Spark
3. Deep Dive into Apache Spark
4. Big Data Analytics with Spark SQL, DataFrames, and Datasets
5. Real-Time Analytics with Spark Streaming and Structured Streaming
6. Notebooks and Dataflows with Spark and Hadoop
7. Machine Learning with Spark and Hadoop
8. Building Recommendation Systems with Spark and Mahout
9. Graph Analytics with GraphX
10. Interactive Analytics with SparkR

## Big Data Analytics By Venkat Ankam Bibliography

- Sales Rank: #1270163 in eBooks
- Published on: 2016-09-28
- Released on: 2016-09-28
- Format: Kindle eBook

 [Download Big Data Analytics ...pdf](#)

 [Read Online Big Data Analytics ...pdf](#)

## **Download and Read Free Online Big Data Analytics By Venkat Ankam**

---

### **Editorial Review**

#### **About the Author**

#### **Venkat Ankam**

Venkat Ankam has over 18 years of IT experience and over 5 years in big data technologies, working with customers to design and develop scalable big data applications. Having worked with multiple clients globally, he has tremendous experience in big data analytics using Hadoop and Spark. He is a Cloudera Certified Hadoop Developer and Administrator and also a Databricks Certified Spark Developer. He is the founder and presenter of a few Hadoop and Spark meetup groups globally and loves to share knowledge with the community. Venkat has delivered hundreds of trainings, presentations, and white papers in the big data sphere. While this is his first attempt at writing a book, many more books are in the pipeline.

### **Users Review**

#### **From reader reviews:**

#### **David Lacey:**

Do you have favorite book? When you have, what is your favorite's book? Book is very important thing for us to be aware of everything in the world. Each publication has different aim or goal; it means that guide has different type. Some people sense enjoy to spend their time to read a book. They are reading whatever they acquire because their hobby will be reading a book. Why not the person who don't like looking at a book? Sometime, particular person feel need book if they found difficult problem or maybe exercise. Well, probably you should have this Big Data Analytics.

#### **Dora Gourley:**

Now a day people who Living in the era exactly where everything reachable by talk with the internet and the resources in it can be true or not call for people to be aware of each information they get. How many people to be smart in acquiring any information nowadays? Of course the solution is reading a book. Examining a book can help folks out of this uncertainty Information especially this Big Data Analytics book because this book offers you rich facts and knowledge. Of course the knowledge in this book hundred percent guarantees there is no doubt in it everybody knows.

#### **Jeffrey David:**

A lot of guide has printed but it differs from the others. You can get it by online on social media. You can choose the most effective book for you, science, witty, novel, or whatever by simply searching from it. It is called of book Big Data Analytics. You can add your knowledge by it. Without leaving behind the printed book, it may add your knowledge and make a person happier to read. It is most essential that, you must aware about reserve. It can bring you from one spot to other place.

**Mary Fix:**

Reading a book make you to get more knowledge from the jawhorse. You can take knowledge and information from your book. Book is composed or printed or outlined from each source this filled update of news. Within this modern era like right now, many ways to get information are available for a person. From media social including newspaper, magazines, science guide, encyclopedia, reference book, new and comic. You can add your knowledge by that book. Are you ready to spend your spare time to spread out your book? Or just in search of the Big Data Analytics when you desired it?

**Download and Read Online Big Data Analytics By Venkat Ankam  
#WZVYL4UXORG**

# **Read Big Data Analytics By Venkat Ankam for online ebook**

Big Data Analytics 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 By Venkat Ankam books to read online.

## **Online Big Data Analytics By Venkat Ankam ebook PDF download**

**Big Data Analytics By Venkat Ankam Doc**

**Big Data Analytics By Venkat Ankam Mobipocket**

**Big Data Analytics By Venkat Ankam EPub**

**WZVYLUORG: Big Data Analytics By Venkat Ankam**