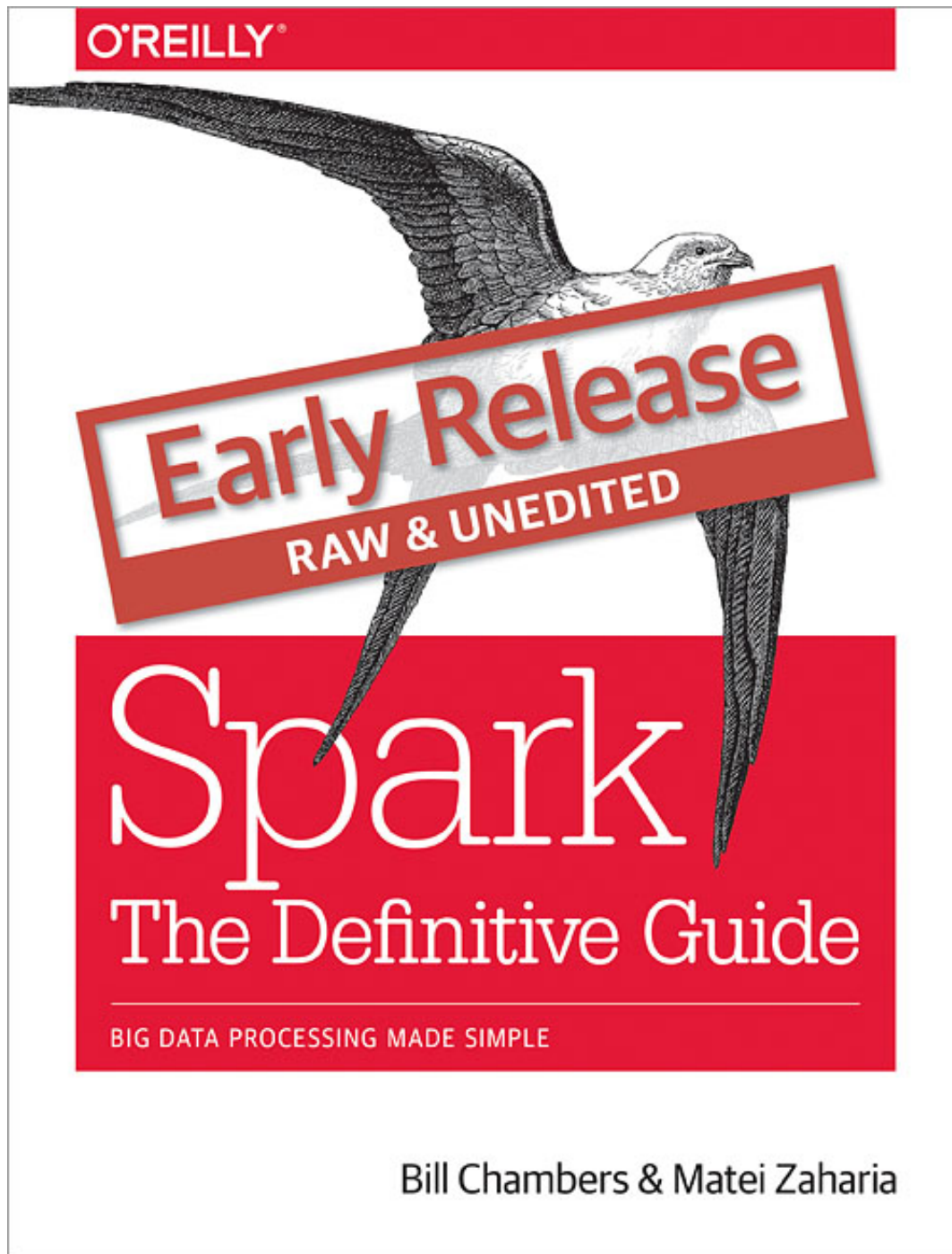


## [\[电子书\]Spark: The Definitive Guide Early Release PDF下载](#)

本书作者：Bill Chambers、Matei Zaharia、Shrey Mehrotra，由O'Reilly Media出版社于2017年1月出版，全书共450页。这里提供的是本书的 Early Release 版本，正式版尚未出版，而且目前还没有完整的内容。由于这本书有Matei Zaharia参与编写，所有很值得一看。

通过本书将学习到以下的知识：

- 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 Spark's Structured Streaming and MLlib for machine learning tasks
- Explore the wider Spark ecosystem, including SparkR and Graph Analysis
- Examine Spark deployment, including coverage of Spark in the Cloud



如果想及时了解Spark、Hadoop或者Hbase相关的文章，欢迎关注微信公共帐号：iteblog\_hadoop

## 本书的章节

- Chapter 1: What is Spark?
- Chapter 2: A Gentle Introduction to Spark
- Chapter 3: Advanced Spark
- Chapter 4: Structured APIs Overview
- Chapter 5: Basic Structured Operations

Chapter 6: Working with Different Types of Data  
Chapter 7: Aggregations  
Chapter 8: Joins  
Chapter 9: Data Sources  
Chapter 10: Spark SQL  
Chapter 11: Datasets  
Chapter 12: Low Level APIs Overview  
Chapter 13: Basic RDD Operations  
Chapter 14: Advanced RDD Operations  
Chapter 15: Distributed Variables  
Chapter 16: How Spark Runs on a Cluster  
Chapter 17: Writing Spark Applications  
Chapter 18: Deploying Spark  
Chapter 19: Monitoring and Debugging  
Chapter 20: Configuring, Optimizing, and Tuning  
Chapter 21: Streaming Concepts  
Chapter 22: What Makes Streaming in Spark Different  
Chapter 23: Structured Streaming  
Chapter 24: Operationalizing Structured Streaming  
Chapter 25: Advanced Analytics and Machine Learning  
Chapter 26: Preprocessing Data  
Chapter 27: Classification  
Chapter 28: Regression  
Chapter 29: Recommendation  
Chapter 30: Clustering  
Chapter 31: Graph Analysis  
Chapter 32: Deep Learning  
Chapter 33: Ecosystem and Community  
Chapter 34: Spark Packages  
Chapter 35: R on Spark

## 下载地址

本博客文章除特别声明，全部都是原创！  
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。  
本文链接：[【】（）](#)