



1st ed., XXIII, 277 p. 64 illus.

### Printed book

Softcover

37,99 € | £32.99 | \$44.99

<sup>[1]</sup>40,65 € (D) | 41,79 € (A) | CHF 45,00

### eBook

29,99 € | £25.99 | \$34.99

<sup>[2]</sup>29,99 € (D) | 29,99 € (A) | CHF 36,00

Available from your library or  
[springer.com/shop](http://springer.com/shop)

### MyCopy <sup>[3]</sup>

Printed eBook for just

€ | \$ 24.99

[springer.com/mycopy](http://springer.com/mycopy)

Mohammed Guller

# Big Data Analytics with Spark

A Practitioner's Guide to Using Spark for Large Scale Data Analysis

- Highlights the role of the Spark in the Big Data Landscape and proves the worth of having strong Spark skills in order to increase subject knowledge and boost your career
- Introduces readers to the concept of functional programming in Scala, one of the core languages supported by Spark
- Key fundamental concepts of Spark are highlighted through practical examples for different scenarios, each showcasing the features of Spark's core and its add-on libraries, Spark SQL, Spark Streaming, GraphX, and MLlib and so on

Big Data Analytics with Spark is a step-by-step guide for learning Spark, which is an open-source fast and general-purpose cluster computing framework for large-scale data analysis. You will learn how to use Spark for different types of big data analytics projects, including batch, interactive, graph, and stream data analysis as well as machine learning. In addition, this book will help you become a much sought-after Spark expert. Spark is one of the hottest Big Data technologies. The amount of data generated today by devices, applications and users is exploding. Therefore, there is a critical need for tools that can analyze large-scale data and unlock value from it. Spark is a powerful technology that meets that need. You can, for example, use Spark to perform low latency computations through the use of efficient caching and iterative algorithms; leverage the features of its shell for easy and interactive Data analysis; employ its fast batch processing and low latency features to process your real time data streams and so on. As a result, adoption of Spark is rapidly growing and is replacing Hadoop MapReduce as the technology of choice for big data analytics. This book provides an introduction to Spark and related big-data technologies. It covers Spark core and its add-on libraries, including Spark SQL, Spark Streaming, GraphX, and MLlib.

Order online at [springer.com](http://springer.com) / or for the Americas call (toll free) 1-800-SPRINGER / or email us at: [customerservice@springernature.com](mailto:customerservice@springernature.com). / For outside the Americas call +49 (0) 6221-345-4301 / or email us at: [customerservice@springernature.com](mailto:customerservice@springernature.com).

The first € price and the £ and \$ price are net prices, subject to local VAT. Prices indicated with [1] include VAT for books; the €(D) includes 7% for Germany, the €(A) includes 10% for Austria. Prices indicated with [2] include VAT for electronic products; 19% for Germany, 20% for Austria. All prices exclusive of carriage charges. Prices and other details are subject to change without notice. All errors and omissions excepted. [3] No discount for MyCopy.

