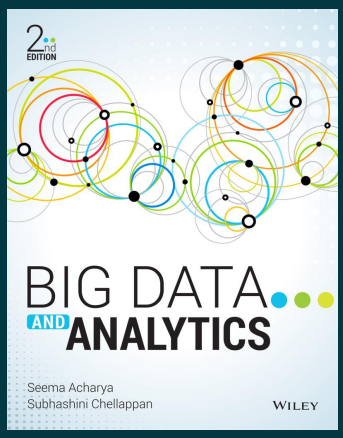
WILEY



Big Data and Analytics, 2ed

By Seema Acharya, Subhashini Chellappan

Paperback 9788126579518 ₹899.00

Pages: 384

Description

BIG DATA is a term used for massive mounds of structured, semi-structured and unstructured data that has the potential to be mined for information. The real power lies not just in having colossal data but in what insights can be drawn from this data to facilitate better and faster decisions. This book Big Data and Analytics is a comprehensive coverage on the concepts and practice of Big Data, Hadoop and Analytics. From the Do It Yourself steps and guidelines to set up a Hadoop Cluster to the deeper understanding of concepts and ample time-tested hands-on practice exercises on the concepts learned, this ONE book has it all!

About the Author

Seema Acharya, Subhashini Chellappan

Seema Acharya is a Senior Lead Principal with the Education, Training and Assessment department of Infosys Limited. She is a technology evangelist, a learning strategist, and an author with over 15+ years of IT experience in learning/education services. She has designed and delivered several large-scale competency development programs across the globe involving organizational competency need analysis, conceptualization, design, development and deployment of competency development programs. She is an educator by choice and vocation

Table of Contents

Chapter 1 Types of Digital Data

What’s in Store?

1.1 Classification of Digital Data

Chapter 2 Introduction to Big Data

What’s in Store?

2.1 Characteristics of Data

2.2 Evolution of Big Data

2.3 Definition of Big Data

2.4 Challenges with Big Data

2.5 What is Big Data?

2.6 Other Characteristics of Data Which are not Definitional Traits of Big Data

2.7 Why Big Data?

2.8 Are We Just an Information Consumer or Do We also Produce Information?

2.9 Traditional Business Intelligence (BI) versus Big Data

2.10 A Typical Data Warehouse Environment
2.11 A Typical Hadoop Environment
2.12 What is New Today?
2.13 What is Changing in the Realms of Big Data?
Chapter 3 Big Data Analytics
What's in Store?
3.1 Where do we Begin?
3.2 What is Big Data Analytics?
3.3 What Big Data Analytics Isn't?
3.4 Why this Sudden Hype Around Big Data Analytics?
3.5 Classification of Analytics
3.6 Greatest Challenges that Prevent Businesses from Capitalizing on Big Data
3.7 Top Challenges Facing Big Data
3.8 Why is Big Data Analytics Important?
3.9 What Kind of Technologies are we Looking Toward to Help Meet the Challenges Posed by Big Data?
3.10 Data Science
3.11 Data Scientist...Your New Best Friend!!!
3.12 Terminologies Used in Big Data Environments
3.13 Basically Available Soft State Eventual Consistency (BASE)
3.14 Few Top Analytics Tools
Chapter 4 The Big Data Technology Landscape
What's in Store?
4.1 NoSQL (Not Only SQL)
4.2 Hadoop
Remind Me
Point Me (Books)
Connect Me (Internet Resources)
Test Me
Chapter 5 Introduction to Hadoop
What's in Store?
5.1 Introducing Hadoop
5.2 Why Hadoop?
5.3 Why not RDBMS?
5.4 RDBMS versus Hadoop
5.5 Distributed Computing Challenges
5.6 History of Hadoop
5.7 Hadoop Overview
5.8 Use Case of Hadoop

5.9 Hadoop Distributors
5.10 HDFS (Hadoop Distributed File System)
5.11 Processing Data with Hadoop
5.12 Managing Resources and Applications with Hadoop YARN (Yet Another Resource Negotiator)
5.13 Interacting with Hadoop Ecosystem
Chapter 6 Introduction to MongoDB
What’s in Store?
6.1 What is MongoDB?
6.2 Why MongoDB?
6.3 Terms Used in RDBMS and MongoDB
6.4 Data Types in MongoDB
6.5 MongoDB Query Language
Chapter 7 Introduction to Cassandra
What’s in Store?
7.1 Apache Cassandra – An Introduction
7.2 Features of Cassandra
7.3 CQL Data Types
7.4 CQLSH
7.5 Keyspaces
7.6 CRUD (Create, Read, Update, and Delete) Operations
7.7 Collections
7.8 Using a Counter
7.9 Time to Live (TTL)
7.10 Alter Commands
7.11 Import and Export
7.12 Querying System Tables
7.13 Practice Examples
Chapter 8 Introduction to MAPREDUCE Programming
What’s in Store?
8.1 Introduction
8.2 Mapper
8.3 Reducer
8.4 Combiner
8.5 Partitioner
8.6 Searching
8.7 Sorting
8.8 Compression
Chapter 9 Introduction to Hive

What’s in Store?
9.1 What is Hive?
9.2 Hive Architecture
9.3 Hive Data Types
9.4 Hive File Format
9.5 Hive Query Language (HQL)
9.6 RCFile Implementation
9.7 SerDe
9.8 User-Defined Function (UDF)
Chapter 10 Introduction to Pig
What’s in Store?
10.1 What is Pig?
10.2 The Anatomy of Pig
10.3 Pig on Hadoop
10.4 Pig Philosophy
10.5 Use Case for Pig: ETL Processing
10.6 Pig Latin Overview
10.7 Data Types in Pig
10.8 Running Pig
10.9 Execution Modes of Pig
10.10 HDFS Commands
10.11 Relational Operators
10.12 Eval Function
10.13 Complex Data Types
10.14 Piggy Bank
10.15 User-Defined Functions (UDF)
10.16 Parameter Substitution
10.17 Diagnostic Operator
10.18 Word Count Example using Pig
10.19 When to use Pig?
10.20 When not to use Pig?
10.21 Pig at Yahoo!
10.22 Pig versus Hive
Chapter 11 JasperReport using Jaspersoft
What’s in Store?
11.1 Introduction to JasperReports
11.2 Connecting to MongoDB NoSQL Database
11.3 Connecting to Cassandra NoSQL Database

Chapter 12 Introduction to Machine Learning

What’s in Store?

12.1 Introduction to Machine Learning

12.2 Machine Learning Algorithms

Chapter 13 Few Interesting Differences

What’s in Store?

13.1 Difference between Data Warehouse and Data Lake

13.2 Difference between RDBMS and HDFS

13.3 Difference between HDFS and HBase

13.4 Hadoop MapReduce versus Pig

13.5 Difference between Hadoop MapReduce and Spark

13.6 Difference between Pig and Hive

Chapter 14 Big Data Trends in 2019 and Beyond

What’s in Store?

14.1 Rise of the New Age “Data Curators”

14.2 CDOs are Stepping Up

14.3 Dark Data in the Cloud

14.4 Streaming the IoT for Machine Learning

14.5 Edge Computing

14.6 Open Source

14.7 Hadoop is Fundamental and will Remain So!

14.8 Chatbots will Get Smarter

14.9 Container(ed) Revolution

14.10 Commoditization of Visualization

Glossary

Index

To purchase this product, please visit:
https://www.wileyindia.com/big-data-and-analytics-2ed.html

Scan to buy