1.	Record Nr.	UNINA9910811945103321
	Autore	Walkowiak Simon
	Titolo	Big data analytics with R : utilize R to uncover hidden patterns in your big data / / Simon Walkowiak
	Pubbl/distr/stampa	Birmingham : , : Packt Publishing, , 2016
	ISBN	1-78646-372-5
	Edizione	[1st edition]
	Descrizione fisica	1 online resource (498 pages) : illustrations
	Collana	Community experience distilled
	Soggetti	R (Computer program language) Data mining Information visualization text and data mining programming language big data cloud computing data processing software
	Lingua di pubblicazione	Inglese
	Formato	Materiale a stampa
	Livello bibliografico	Monografia
	Note generali	Includes index.
	Sommario/riassunto	Utilize R to uncover hidden patterns in your Big Data About This Book Perform computational analyses on Big Data to generate meaningful results Get a practical knowledge of R programming language while working on Big Data platforms like Hadoop, Spark, H2O and SQL/NoSQL databases, Explore fast, streaming, and scalable data analysis with the most cutting-edge technologies in the market Who This Book Is For This book is intended for Data Analysts, Scientists, Data Engineers, Statisticians, Researchers, who want to integrate R with their current or future Big Data workflows. It is assumed that readers have some experience in data analysis and understanding of data management and algorithmic processing of large quantities of data, however they may lack specific skills related to R. What You Will Learn Learn about current state of Big Data processing using R programming

language and its powerful statistical capabilities Deploy Big Data analytics platforms with selected Big Data tools supported by R in a cost-effective and time-saving manner Apply the R language to real-world Big Data problems on a multi-node Hadoop cluster, e.g. electricity consumption across various socio-demographic indicators and bike share scheme usage Explore the compatibility of R with Hadoop, Spark, SQL and NoSQL databases, and H2O platform In Detail Big Data analytics is the process of examining large and complex data sets that often exceed the computational capabilities. R is a leading programming language of data science, consisting of powerful functions to tackle all problems related to Big Data processing. The book will begin with a brief introduction to the Big Data world and its current industry standards. With introduction to the R language and presenting its development, structure, applications in real world, and its shortcomings. Book will progress towards revision of major R functions for data management and transformations. Readers will be introduce to Cloud based Big Data solutions (e.g. Amazon EC2 instances and Amazon RDS, Microsoft Azure and its HDInsight clusters) and also provide guidance on R connectivity with relational and non-relational databases such as MongoDB and HBase etc. It will further expand to include Big Data tools such as Apache Hadoop ecosystem, HDFS and MapReduce frameworks. Also other R compatible tools such as Apache Spark, its machine learning library Spark MLlib, as well as H2O. Style and approach This book will serve as a practical guide...