1. Record Nr. UNINA9910254561803321

Autore	Mailund Thomas
Titolo	Beginning Data Science in R : Data Analysis, Visualization, and Modelling for the Data Scientist / / by Thomas Mailund
Pubbl/distr/stampa	Berkeley, CA : , : Apress : , : Imprint : Apress, , 2017
Edizione	[1st ed. 2017.]
Descrizione fisica	1 online resource (XXVII, 352 p. 100 illus.)
Disciplina	006.312
Soggetti	Big data Programming languages (Electronic computers) R (Computer program language) Big Data Programming Languages, Compilers, Interpreters
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	1. Introduction to R programming -- 2. Reproducible analysis -- 3. Data manipulation -- 4. Visualizing and exploring data -- 5. Working with large data sets -- 6. Supervised learning -- 7. Unsupervised learning -- 8. More R programming -- 9. Advanced R programming -- 10. Object oriented programming -- 11. Building an R package -- 12. Testing and checking -- 13. Version control -- 14. Profiling and optimizing.
Sommario/riassunto	Discover best practices for data analysis and software development in R and start on the path to becoming a fully-fledged data scientist. This book teaches you techniques for both data manipulation and visualization and shows you the best way for developing new software packages for R. Data Science in R details how data science is a combination of statistics, computational science, and machine learning. You'll see how to efficiently structure and mine data to extract useful patterns and build mathematical models. This requires computational methods and programming, and R is an ideal programming language for this. This book is based on a number of lecture notes for classes the author has taught on data science and statistical programming using the R programming language. Modern data analysis requires

computational skills and usually a minimum of programming. You will: Perform data science and analytics using statistics and the R programming language Visualize and explore data, including working with large data sets found in big data Build an R package Test and check your code Practice version control Profile and optimize your code.