1. Record Nr. UNINA9910299980203321

Autore	Nolan Deborah
Titolo	XML and Web Technologies for Data Sciences with R / / by Deborah Nolan, Duncan Temple Lang
Pubbl/distr/stampa	New York, NY : , : Springer New York : , : Imprint : Springer, , 2014
ISBN	1-4614-7900-2
Edizione	[1st ed. 2014.]
Descrizione fisica	1 online resource (677 p.)
Collana	Use R!, , 2197-5736
Disciplina	004.678
Soggetti	Statistics Programming languages (Electronic computers) R (Computer program language) Statistics and Computing/Statistics Programs Programming Languages, Compilers, Interpreters Statistics, general
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Data Formats XML and JSON -- Web Technologies, Getting Data from the Web -- General XML Application Areas -- Bibliography -- General Index -- R Function and Parameter Index -- R Package Index -- R Class Index -- Colophon.
Sommario/riassunto	Web technologies are increasingly relevant to scientists working with data, for both accessing data and creating rich dynamic and interactive displays. The XML and JSON data formats are widely used in Web services, regular Web pages and JavaScript code, and visualization formats such as SVG and KML for Google Earth and Google Maps. In addition, scientists use HTTP and other network protocols to scrape data from Web pages, access REST and SOAP Web Services, and interact with NoSQL databases and text search applications. This book provides a practical hands-on introduction to these technologies, including high-level functions the authors have developed for data scientists. It describes strategies and approaches for extracting data from HTML, XML, and JSON formats and how to programmatically access data from the Web. Along with these general skills, the authors illustrate several applications that are relevant to data scientists, such as reading and

writing spreadsheet documents both locally and via GoogleDocs, creating interactive and dynamic visualizations, displaying spatial-temporal displays with Google Earth, and generating code from descriptions of data structures to read and write data. These topics demonstrate the rich possibilities and opportunities to do new things with these modern technologies. The book contains many examples and case-studies that readers can use directly and adapt to their own work. The authors have focused on the integration of these technologies with the R statistical computing environment. However, the ideas and skills presented here are more general, and statisticians who use other computing environments will also find them relevant to their work. Deborah Nolan is Professor of Statistics at University of California, Berkeley. Duncan Temple Lang is Associate Professor of Statistics at University of California, Davis and has been a member of both the S and R development teams.