1.	Record Nr.	UNISA996418265903316

Autore Jockers Matthew L

Titolo Text Analysis with R [[electronic resource]] : For Students of Literature / / by Matthew L. Jockers, Rosamond Thalken

Pubbl/distr/stampa Cham : , : Springer International Publishing : , : Imprint : Springer, , 2020

ISBN 3-030-39643-6

Edizione [2nd ed. 2020.]

Descrizione fisica 1 online resource (XXIV, 276 p. 33 illus., 12 illus. in color.)

Collana Quantitative Methods in the Humanities and Social Sciences, , 2199-0956

Disciplina 410.285

Soggetti Statistics
Computational linguistics
Humanities—Digital libraries
Technology in literature
Application software
R (Computer program language)
Statistics and Computing/Statistics Programs
Computational Linguistics
Statistics for Social Sciences, Humanities, Law
Digital Humanities
Literature and Technology/Media
Computer Appl. in Arts and Humanities

Lingua di pubblicazione Inglese

Formato Materiale a stampa

Livello bibliografico Monografia

Nota di contenuto Part I Microanalysis -- 1 R Basics -- 2 First Foray into Text Analysis with R -- 3 Accessing and Comparing Word Frequency Data -- 4 Token Distribution and Regular Expressions -- 5 Token Distribution Analysis by Chapter -- 6 Correlation -- 7 Measures of Lexical Variety -- 8 Hapax Richness -- 9 Do it KWIC -- 10 Do it KWIC(er) (And Better) -- Part II Metadata -- 11 Introduction to dplyr -- 12 Parsing TEI XML- 13 Parsing and Analyzing Hamlet -- 14 Sentiment Analysis -- Part III Macroanalysis -- 15 Clustering -- 16 Classification -- 17 Topic Modeling -- 18 Part of Speech Tagging and Named Entity Recognition -- Appendices -- Index -- List of Tables -- List of Figures.

Sommario/riassunto

Now in its second edition, Text Analysis with R provides a practical introduction to computational text analysis using the open source programming language R. R is an extremely popular programming language, used throughout the sciences; due to its accessibility, R is now used increasingly in other research areas. In this volume, readers immediately begin working with text, and each chapter examines a new technique or process, allowing readers to obtain a broad exposure to core R procedures and a fundamental understanding of the possibilities of computational text analysis at both the micro and the macro scale. Each chapter builds on its predecessor as readers move from small scale "microanalysis" of single texts to large scale "macroanalysis" of text corpora, and each concludes with a set of practice exercises that reinforce and expand upon the chapter lessons. The book's focus is on making the technical palatable and making the technical useful and immediately gratifying. Text Analysis with R is written with students and scholars of literature in mind but will be applicable to other humanists and social scientists wishing to extend their methodological toolkit to include quantitative and computational approaches to the study of text. Computation provides access to information in text that readers simply cannot gather using traditional qualitative methods of close reading and human synthesis. This new edition features two new chapters: one that introduces dplyr and tidyr in the context of parsing and analyzing dramatic texts to extract speaker and receiver data, and one on sentiment analysis using the syuzhet package. It is also filled with updated material in every chapter to integrate new developments in the field, current practices in R style, and the use of more efficient algorithms.