

1. Record Nr.	UNINA9910464122103321
Autore	Toomey Dan
Titolo	R for data science : learn and explore the fundamentals of data science with R // Dan Toomey
Pubbl/distr/stampa	Birmingham, England : , : Packt Publishing, , 2014 ©2014
ISBN	1-78439-265-0
Descrizione fisica	1 online resource (364 p.)
Collana	Community Experience Distilled
Disciplina	519.502855133
Soggetti	R (Computer program language) Mathematical statistics - Data processing Electronic books.
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di contenuto	Cover; Copyright; Credits; About the Author; About the Reviewers; www.PacktPub.com; Table of Contents; Preface; Chapter 1: Data Mining Patterns; Cluster analysis; K-means clustering; Usage; Example; K-medoids clustering; Usage; Example; Hierarchical clustering; Usage; Example; Expectation-maximization; Usage; List of model names; Example; Density estimation; Usage; Example; Anomaly detection; Show outliers; Example; Example; Another anomaly detection example; Calculating anomalies; Usage; Example 1; Example 2; Association rules; Mine for associations; Usage; Example; Questions; Summary Chapter 2: Data Mining SequencesPatterns; Eclat; Usage; Using eclat to find similarities in adult behavior; Finding frequent items in a dataset; An example focusing on highest frequency; arulesNBMiner; Usage; Mining the Agrawal data for frequent sets; Apriori; Usage; Evaluating associations in a shopping basket; Determining sequences using TraMineR; Usage; Determining sequences in training and careers; Similarities in the sequence; Sequence metrics; Usage; Example; Questions; Summary; Chapter 3: Text Mining; Packages; Text processing; Example; Creating a corpus; Text clusters; Word graphics Analyzing the XML textQuestions; Summary; Chapter 4: Data Analysis - Regression Analysis; Packages; Simple regression; Multiple regression; Multivariate regression analysis; Robust regression; Questions;

Summary; Chapter 5: Data Analysis - Correlation; Packages; Correlation; Example; Visualizing correlations; Covariance; Pearson correlation; Polychoric correlation; Tetrachoric correlation; A heterogeneous correlation matrix; Partial correlation; Questions; Summary; Chapter 6: Data Analysis - Clustering; Packages; K-means clustering; Example; Optimal number of clusters; Medoids clusters The cascadeKM function Selecting clusters based on Bayesian information; Affinity propagation clustering; Gap statistic to estimate the number of clusters; Hierarchical clustering; Questions; Summary; Chapter 7: Data Visualization - R Graphics; Packages; Interactive graphics; The latticist package; Bivariate binning display; Mapping; Plotting points on a map; Plotting points on a world map; Google Maps; The ggplot2 package; Questions; Summary; Chapter 8: Data Visualization - Plotting; Packages; Scatter plots; Regression line; A loess line; scatterplot; Scatterplot matrices splom - display matrix data; pairs - plot matrix data; Density scatter plots; Bar charts and plots; Bar plot; Usage; Bar chart; ggplot2; Word cloud; Questions; Summary; Chapter 9: Data Visualization - 3D; Packages; Generating 3D graphics; Lattice Cloud - 3D scatterplot; scatterplot3d; scatter3d; cloud3d; RgoogleMaps; vrmIgenbar3D; Big Data; pbdR; bigmemory; Research areas; Rcpp; parallel; microbenchmark; pqR; SAP integration; roxygen2; bioconductor; swirl; pipes; Questions; Summary; Chapter 10: Machine Learning in Action; Packages; Dataset; Data partitioning; Model; Linear model; Prediction Logistic regression

Sommario/riassunto

If you are a data analyst who has a firm grip on some advanced data analysis techniques and wants to learn how to leverage the features of R, this is the book for you. You should have some basic knowledge of the R language and should know about some data science topics.
