

Springer

2nd
edition2nd ed. 2019, XV, 524 p.
187 illus., 80 illus. in color.**Printed book**

Softcover

Printed book

Softcover

ISBN 978-3-030-00270-1

\$ 74,99

Available

Discount group

Professional Books (2)

Product category

Graduate/advanced undergraduate textbook

Series

Undergraduate Topics in Computer Science

Computer Science : Data Structures

Mirkin, Boris, National Research University Higher School of Economics, Moscow, Russia

Core Data Analysis: Summarization, Correlation, and Visualization

- Focuses on the encoder-decoder interpretation of summarization methods, such as Principal Component Analysis and K-means clustering
- Supplies an in-depth description of K-means partitioning including a data-driven mathematical theory
- Covers novel topics such as Google PageRank ranking and Consensus clustering as interlaced within the general framework
- Includes a multitude of worked examples, case studies and questions (with answers)

This text examines the goals of data analysis with respect to enhancing knowledge, and identifies data summarization and correlation analysis as the core issues. Data summarization, both quantitative and categorical, is treated within the encoder-decoder paradigm bringing forward a number of mathematically supported insights into the methods and relations between them. Two Chapters describe methods for categorical summarization: partitioning, divisive clustering and separate cluster finding and another explain the methods for quantitative summarization, Principal Component Analysis and PageRank. Features: · An in-depth presentation of K-means partitioning including a corresponding Pythagorean decomposition of the data scatter. · Advice regarding such issues as clustering of categorical and mixed scale data, similarity and network data, interpretation aids, anomalous clusters, the number of clusters, etc. · Thorough attention to data-driven modelling including a number of mathematically stated relations between statistical and geometrical concepts including those between goodness-of-fit criteria for decision trees and data standardization, similarity and consensus clustering, modularity clustering and uniform partitioning.

Order online at [springer.com/booksellers](https://www.springer.com/booksellers)

Springer Nature Customer Service Center LLC

233 Spring Street

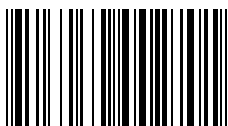
New York, NY 10013

USA

T: +1-800-SPRINGER NATURE

(777-4643) or 212-460-1500

customerservice@springernature.com



ISBN 978-3-030-00270-1 / BIC: UMB / SPRINGER NATURE: SCI15017

Prices and other details are subject to change without notice. All errors and omissions excepted. Americas: Tax will be added where applicable. Canadian residents please add PST, QST or GST. Please add \$5.00 for shipping one book and \$ 1.00 for each additional book. Outside the US and Canada add \$ 10.00 for first book, \$5.00 for each additional book. If an order cannot be fulfilled within 90 days, payment will be refunded upon request. Prices are payable in US currency or its equivalent.

Part of **SPRINGER NATURE**