



2nd ed. 2019, XV, 524 p. 187 illus., 80 illus. in color.

Printed book

Softcover

59,99 € | £54.99 | \$74.99

^[1]64,19 € (D) | 65,99 € (A) | CHF 71,00

eBook

51,16 € | £43.99 | \$59.99

^[2]51,16 € (D) | 51,16 € (A) | CHF 56,50

Available from your library or springer.com/shop

MyCopy ^[3]

Printed eBook for just

€ | \$ 24.99

springer.com/mycopy

Boris Mirkin

Core Data Analysis: Summarization, Correlation, and Visualization

Series: Undergraduate Topics in Computer Science

- 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 / or for the Americas call (toll free) 1-800-SPRINGER / or email us at: customerservice@springernature.com. / For outside the Americas call +49 (0) 6221-345-4301 / or email us at: customerservice@springernature.com.

The first € price and the £ and \$ price are net prices, subject to local VAT. Prices indicated with [1] include VAT for books; the €(D) includes 7% for Germany, the €(A) includes 10% for Austria. Prices indicated with [2] include VAT for electronic products; 19% for Germany, 20% for Austria. All prices exclusive of carriage charges. Prices and other details are subject to change without notice. All errors and omissions excepted. [3] No discount for MyCopy.

