| | | |
|---|---|---|
| 1. | Record Nr. | UNINA9911019310003321 |
| | Autore | Billard L (Lynne), <1943-> |
| | Titolo | Symbolic data analysis : conceptual statistics and data mining / / Lynne Billard, Edwin Diday |
| | Pubbl/distr/stampa | Chichester, England ; ; Hoboken, NJ, : John Wiley & Sons Inc., c2006 |
| | ISBN | 9786613622143 |
| | | 9781280592317 |
| | | 1280592311 |
| | | 9780470090183 |
| | | 0470090189 |
| | | 9780470090176 |
| | | 0470090170 |
| | Descrizione fisica | 1 online resource (331 p.) |
| | Collana | Wiley series in computational statistics |
| | Altri autori (Persone) | DidayE |
| | Disciplina | 519.5/35 |
| | | 519.535 |
| | Soggetti | Data mining |
| | | Multivariate analysis |
| | 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 indexes. |
| | Nota di contenuto | Symbolic Data Analysis; Contents; 1 Introduction; References; 2 Symbolic Data; 2.1 Symbolic and Classical Data; 2.1.1 Types of data; 2.1.2 Dependencies in the data; 2.2 Categories, Concepts, and Symbolic Objects; 2.2.1 Preliminaries; 2.2.2 Descriptions, assertions, extents; 2.2.3 Concepts of concepts; 2.2.4 Some philosophical aspects; 2.2.5 Fuzzy, imprecise, and conjunctive data; 2.3 Comparison of Symbolic and Classical Analyses; Exercises; References; 3 Basic Descriptive Statistics: One Variate; 3.1 Some Preliminaries; 3.2 Multi-Valued Variables; 3.3 Interval-Valued Variables 3.4 Modal Multi-Valued Variables3.5 Modal Interval-Valued Variables; Exercises; References; 4 Descriptive Statistics: Two or More Variates; 4.1 Multi-Valued Variables; 4.2 Interval-Valued Variables; 4.3 Modal Multi-Valued Variables; 4.4 Modal Interval-Valued Variables; 4.5 Baseball Interval-Valued Dataset; 4.5.1 The data: actual and virtual datasets; 4.5.2 Joint histograms; 4.5.3 Guiding principles; 4.6 Measures |

| | |
|---|---|
| Sommario/riassunto | With the advent of computers, very large datasets have become routine. Standard statistical methods don't have the power or flexibility to analyse these efficiently, and extract the required knowledge. An alternative approach is to summarize a large dataset in such a way that the resulting summary dataset is of a manageable size and yet retains as much of the knowledge in the original dataset as possible. One consequence of this is that the data may no longer be formatted as single values, but be represented by lists, intervals, distributions, etc. The summarized data have their own internal s |