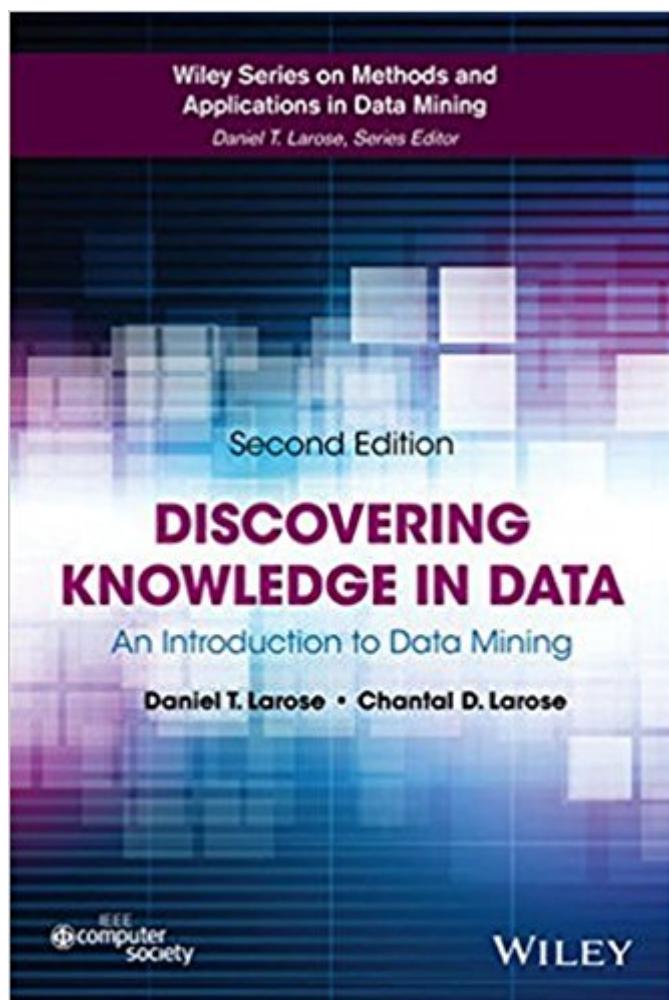


The book was found

# Discovering Knowledge In Data: An Introduction To Data Mining (Wiley Series On Methods And Applications In Data Mining)



## Synopsis

NOTE: Solution manual is available in the companion site. The link is provided on page 19 in the second paragraph of section 2.3 in the book. The field of data mining lies at the confluence of predictive analytics, statistical analysis, and business intelligence. Due to the ever-increasing complexity and size of data sets and the wide range of applications in computer science, business, and health care, the process of discovering knowledge in data is more relevant than ever before. This book provides the tools needed to thrive in today's big data world. The author demonstrates how to leverage a company's existing databases to increase profits and market share, and carefully explains the most current data science methods and techniques. The reader will learn data mining by doing data mining. By adding chapters on data modelling preparation, imputation of missing data, and multivariate statistical analysis, *Discovering Knowledge in Data*, Second Edition remains the eminent reference on data mining. The second edition of a highly praised, successful reference on data mining, with thorough coverage of big data applications, predictive analytics, and statistical analysis. Includes new chapters on Multivariate Statistics, Preparing to Model the Data, and Imputation of Missing Data, and an Appendix on Data Summarization and Visualization Offers extensive coverage of the R statistical programming language Contains 280 end-of-chapter exercises Includes a companion website with further resources for all readers, and Powerpoint slides, a solutions manual, and suggested projects for instructors who adopt the book

## Book Information

Series: Wiley Series on Methods and Applications in Data Mining (Book 4)

Hardcover: 336 pages

Publisher: Wiley; 2 edition (July 8, 2014)

Language: English

ISBN-10: 0470908742

ISBN-13: 978-0470908747

Product Dimensions: 6.4 x 1 x 9.6 inches

Shipping Weight: 4 ounces (View shipping rates and policies)

Average Customer Review: 4.2 out of 5 stars 12 customer reviews

Best Sellers Rank: #74,611 in Books (See Top 100 in Books) #17 in Books > Computers & Technology > Databases & Big Data > Data Warehousing #49 in Books > Computers & Technology > Databases & Big Data > Data Mining #96 in Books > Textbooks > Computer Science > Database Storage & Design

## Customer Reviews

NOTE: Solution manual is available in the companion site. The link is provided on page 19 in the second paragraph of section 2.3 The field of data mining lies at the confluence of predictive analytics, statistical analysis, and business intelligence. Due to the ever-increasing complexity and size of data sets and the wide range of applications in computer science, business, and health care, the process of discovering knowledge in data is more relevant than ever before. This book provides the tools needed to thrive in todayâ™s big data world. The author demonstrates how to leverage a companyâ™s existing databases to increase profits and market share, and carefully explains the most current data science methods and techniques. The reader will âœlearn data mining by doing data miningâ•. By adding chapters on data modelling preparation, imputation of missing data, and multivariate statistical analysis, Discovering Knowledge in Data, Second Edition remains the eminent reference on data mining.

Daniel T. Larose earned his PhD in Statistics at the University of Connecticut. He is Professor of Mathematical Sciences and Director of the Data Mining programs at Central Connecticut State University. His consulting clients have included Microsoft, Forbes Magazine, the CIT Group, KPMG International, Computer Associates, and Deloitte, Inc. This is Laroseâ™s fourth book for Wiley. Chantal D. Larose is a PhD candidate in Statistics at the University of Connecticut. Her research focuses on the imputation of missing data and model-based clustering. She has taught undergraduate statistics since 2011, and has done statistical consulting for DataMiningConsultant.com, LLC.

What a great introduction to Data Science! I just completed my undergrad in Applied Mathematics and am beginning a graduate degree in Data Mining. I bought this book because it is necessary for my program. I am surprised at how much I enjoy reading this book. First of all, it is written in a way that informs while not overwhelming. There are plenty of examples and graphics throughout the book that add up to make for a walkthrough more than a talking to. I also like that Dr. Larose is a Statistician. He offers insights that can only be gained from understanding the underlying mathematics behind Data Science. I find that as I read through it, and I work through the R exercises and the Hands On Data exercises, I am becoming more confident and optimistic about my future as a Data Scientist. Of course, further study is necessary, but Discovering Knowledge in Data:... is an excellent beginning.

There is a lot to like about this book, but it has some unfortunate flaws. Note that it is part of a Data Mining trilogy. The other two books are:Â Data Mining Methods and ModelsÂ andÂ Data Mining the Web: Uncovering Patterns in Web Content, Structure, and Usage. My initial reaction was more negative as I feel strongly about the issues that this book addresses poorly. However, I find myself turning to this book again and again. I would endorse it highly, but with a caution or two. The very best features of the book are the exceptionally clear explanations of complicated algorithms. In particular, Chapters 6 and 7 and their explanations of Decision Trees and Neural Nets are just perfect for both new and veteran analysts who want to understand what is happening "under the hood". Those chapters are stand-outs, but all of the 80%+ part of the book that describes algorithms in detail (clear, careful, and readable detail) is uniformly excellent. For some readers, it may be the first time that the techniques really make sense to them. Now the flaws. The three book format is, frankly, annoying. The second book and third books are much weaker, but the it was clearly designed as a trilogy, so it is hard to recommend the first to a client without at least implicitly recommending the second. Spending my reading time well is more important to me than my reading budget, but the set of three costs more than \$200. Unless you plan on an entire shelf of related books, like me, I can't recommend the entire set. The other flaw is less obvious, and is the one that concerns me the most. Although this book cites Dorian Pyle's excellent book ... it seems to miss the whole point. Data Mining data prep is quite different from data prep for statistics. Although the two areas share a lot in common, and while mastery of statistics is a good thing for data miners, this is one of the differences between the two disciplines. Data cleaning and data reduction are critical, but this book suggests that this is accomplished by the human doing all possible bivariate.

Recommendations of factor analysis and log transformations abound, but never with cautions of when that is unnecessary or even a bad idea - something Pyle's book explores. Also, transformations like binning come off as something the analyst does during data exploration, getting it perfect before modeling. Sounds like statistics data prep to me - not data mining data prep. If anyone has ever completed data prep without preliminary modeling, or has modeled without having to revisit data prep, I have never heard of it. If a novice data miner were to take the advice too literally, they could get themselves into trouble. This would be especially true of a reader that is well versed in statistics - there is a predictable set of mistakes awaiting the classically trained on their first data mining project! My advice? There is a lot to benefit from here. All of the "white box" walk through examples are great. Consider buying this book, the Pyle bookÂ Data Preparation for Data Mining (The Morgan Kaufmann Series in Data Management Systems), and Berry and LinoffÂ Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, while skipping

the other two in this trilogy. Use this book for the algorithm explanations, but be cautious otherwise. The screen shots and discussion of Clementine may be helpful to you, but note that Clementine 8.5 was used.

Not the best book on the topic, but very well done and concise. I liked the writing style and that it wasn't a huge 500 pager.

Very useful for introductory class in Data Mining!

The book explain the K-nearest neighborhood, Decision tree algorithm (C 4.5 and CART), neural network, K-means and Hierarchical Algorithms in a very efficient way. Also it explain the -processing of data and the steps to implement data mining project. It could longer and cover more materials.

Book was correct and shipped fast

I had it for an grad school data mining course. The subject matter wasn't all that hard, but this book made it seem that way. The examples only made my head spin. I learned the material from the web. The book was unnecessary.

Good

[Download to continue reading...](#)

Discovering Knowledge in Data: An Introduction to Data Mining (Wiley Series on Methods and Applications in Data Mining) Big Data For Business: Your Comprehensive Guide to Understand Data Science, Data Analytics and Data Mining to Boost More Growth and Improve Business - Data Analytics Book, Series 2 Data Classification: Algorithms and Applications (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series) Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data Book 1) Data Analytics: Applicable Data Analysis to Advance Any Business Using the Power of Data Driven Analytics (Big Data Analytics, Data Science, Business Intelligence Book 6) Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness (Chandos

Knowledge Management) Building Winning Algorithmic Trading Systems, + Website: A Trader's Journey From Data Mining to Monte Carlo Simulation to Live Trading (Wiley Trading) Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data Functional Polymer Coatings: Principles, Methods, and Applications (Wiley Series on Polymer Engineering and Technology) Construction Dewatering: New Methods and Applications (Wiley Series of Practical Construction Guides) Data Mining: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data (Princeton Series in Modern Observational Astronomy) Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems) Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) Data Mining: Practical Machine Learning Tools and Techniques, Third Edition (Morgan Kaufmann Series in Data Management Systems) Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems) Data Mining for Business Analytics: Concepts, Techniques, and Applications with JMP Pro Data Mining for Business Analytics: Concepts, Techniques, and Applications with XLMiner

[Contact Us](#)

[DMCA](#)

[Privacy](#)

[FAQ & Help](#)