

Foundations Of Multidimensional And Metric Data Structures The Morgan Kaufmann Series In Computer Graphics

This practical, applications-oriented book describes essential tools for efficiently handling massive amounts of data. This book provides a systematic development of tensor methods in statistics, beginning with the study of multivariate moments and cumulants. The effect on moment arrays and on cumulant arrays of making linear or affine transformations of the variables is studied. Because of their importance in statistical theory, invariant functions of the cumulants are studied in some detail. This is followed by an examination of the effect of making a polynomial transformation of the original variables. The fundamental operation of summing over complementary set partitions is introduced at this stage. This operation shapes the notation and pervades much of the remainder of the book. The necessary lattice-theory is discussed and suitable tables of complementary set partitions are provided. Subsequent chapters deal with asymptotic approximations based on Edgeworth expansion and saddlepoint expansion. The saddlepoint expansion is introduced via the Legendre transformation of the cumulant generating function, also known as the conjugate function of the cumulant generating function. A recurring theme is that, with suitably chosen notation, multivariate calculations are often simpler and more transparent than the corresponding univariate calculations. The final two chapters deal with likelihood ratio statistics, maximum likelihood estimation and the effect on inferences of conditioning on ancillary or approximately ancillary statistics. The Bartlett adjustment factor is derived in the general case and simplified for certain types of generalized linear models. Finally, Barndorff-Nielsen's formula for the conditional distribution of the maximum likelihood estimator is derived and discussed. More than 200 Exercises are provided to illustrate the uses of tensor methodology.

Why bibliometrics is useful for understanding the global dynamics of science but generate perverse effects when applied inappropriately in research evaluation and university rankings. The research evaluation market is booming. "Ranking," "metrics," "h-index," and "impact factors" are reigning buzzwords. Government and research administrators want to evaluate everything—teachers, professors, training programs, universities—using quantitative indicators. Among the tools used to measure "research excellence," bibliometrics—aggregate data on publications and citations—has become dominant. Bibliometrics is hailed as an "objective" measure of research quality, a quantitative measure more useful than "subjective" and intuitive evaluation methods such as peer review that have been used since scientific papers were first published in the seventeenth century. In this book, Yves Gingras offers a spirited argument against an unquestioning reliance on bibliometrics as an indicator of research quality. Gingras shows that bibliometric rankings have no real scientific validity, rarely measuring what they pretend to. Although the study of publication and citation patterns, at the proper scales, can yield insights on the global dynamics of science over time, ill-defined quantitative indicators often generate perverse and unintended effects on the direction of research. Moreover, abuse of bibliometrics occurs when data is manipulated to boost rankings. Gingras looks at the politics of evaluation and argues that using numbers can be a way to control scientists and diminish their autonomy in the evaluation process. Proposing precise criteria for establishing the validity of indicators at a given scale of analysis, Gingras questions why universities are so eager to let invalid indicators influence their research strategy.

Winner of the Shingo Publication Award Accelerate your organization to win in the marketplace. How can we apply technology to drive business value? For years, we've been told that the performance of software delivery teams doesn't matter—that it can't provide a competitive advantage to our companies. Through four years of groundbreaking research to include data collected from the State of DevOps reports conducted with Puppet, Dr. Nicole Forsgren, Jez Humble, and Gene Kim set out to find a way to measure software delivery performance—and what drives it—using rigorous statistical methods. This book presents both the findings and the science behind that research, making the information accessible for readers to apply in their own organizations. Readers will discover how to measure the performance of their teams, and what capabilities they should invest in to drive higher performance. This book is ideal for management at every level.

Class-tested and coherent, this textbook teaches classical and web information retrieval, including web search and the related areas of text classification and text clustering from basic concepts. It gives an up-to-date treatment of all aspects of the design and implementation of systems for gathering, indexing, and searching documents; methods for evaluating systems; and an introduction to the use of machine learning methods on text collections. All the important ideas are explained using examples and figures, making it perfect for introductory courses in information retrieval for advanced undergraduates and graduate students in computer science. Based on feedback from extensive classroom experience, the book has been carefully structured in order to make teaching more natural and effective. Slides and additional exercises (with solutions for lecturers) are also available through the book's supporting website to help course instructors prepare their lectures.

Neurofeedback is utilized by over 10,000 clinicians worldwide with new techniques and uses being found regularly. Z Score Neurofeedback is a new technique using a normative database to identify and target a specific individual's area of dysregulation allowing for faster and more effective treatment. The book describes how to perform z Score Neurofeedback, as well as research indicating its effectiveness for a variety of disorders including pain, depression, anxiety, substance abuse, PTSD, ADHD, TBI, headache, frontal lobe disorders, or for cognitive enhancement. Suitable for clinicians as well as researchers this book is a one stop shop for those looking to understand and use this new technique. Contains protocols to implement Z score neurofeedback Reviews research on disorders for which this is effective treatment Describes advanced techniques and applications

This book describes established and advanced methods for reducing the dimensionality of numerical databases. Each description starts from intuitive ideas, develops the necessary mathematical details, and ends by outlining the algorithmic implementation. The text provides a lucid summary of facts and concepts relating to well-known methods as well as recent developments in nonlinear dimensionality reduction. Methods are all described from a unifying point of view, which helps to highlight their respective strengths and shortcomings. The presentation will appeal to statisticians, computer scientists and data analysts, and other practitioners having a basic background in statistics or computational learning.

A self-contained introduction to the fundamentals of mathematical analysis Mathematical Analysis: A Concise Introduction presents the foundations of analysis and illustrates its role in mathematics. By focusing on the essentials, reinforcing learning through exercises, and featuring a unique "learn by doing" approach, the book develops the reader's proof writing skills and establishes fundamental comprehension of analysis that is essential for further exploration of pure and applied mathematics. This book is directly applicable to areas such as differential equations, probability theory, numerical analysis, differential geometry, and

functional analysis. Mathematical Analysis is composed of three parts: Part One presents the analysis of functions of one variable, including sequences, continuity, differentiation, Riemann integration, series, and the Lebesgue integral. A detailed explanation of proof writing is provided with specific attention devoted to standard proof techniques. To facilitate an efficient transition to more abstract settings, the results for single variable functions are proved using methods that translate to metric spaces. Part Two explores the more abstract counterparts of the concepts outlined earlier in the text. The reader is introduced to the fundamental spaces of analysis, including L_p spaces, and the book successfully details how appropriate definitions of integration, continuity, and differentiation lead to a powerful and widely applicable foundation for further study of applied mathematics. The interrelation between measure theory, topology, and differentiation is then examined in the proof of the Multidimensional Substitution Formula. Further areas of coverage in this section include manifolds, Stokes' Theorem, Hilbert spaces, the convergence of Fourier series, and Riesz' Representation Theorem. Part Three provides an overview of the motivations for analysis as well as its applications in various subjects. A special focus on ordinary and partial differential equations presents some theoretical and practical challenges that exist in these areas. Topical coverage includes Navier-Stokes equations and the finite element method. Mathematical Analysis: A Concise Introduction includes an extensive index and over 900 exercises ranging in level of difficulty, from conceptual questions and adaptations of proofs to proofs with and without hints. These opportunities for reinforcement, along with the overall concise and well-organized treatment of analysis, make this book essential for readers in upper-undergraduate or beginning graduate mathematics courses who would like to build a solid foundation in analysis for further work in all analysis-based branches of mathematics.

Multidimensional scaling (MDS) is a technique for the analysis of similarity or dissimilarity data on a set of objects. Such data may be intercorrelations of test items, ratings of similarity on political candidates, or trade indices for a set of countries. MDS attempts to model such data as distances among points in a geometric space. The main reason for doing this is that one wants a graphical display of the structure of the data, one that is much easier to understand than an array of numbers and, moreover, one that displays the essential information in the data, smoothing out noise. There are numerous varieties of MDS. Some facets for distinguishing among them are the particular type of geometry into which one wants to map the data, the mapping function, the algorithms used to find an optimal data representation, the treatment of statistical error in the models, or the possibility to represent not just one but several similarity matrices at the same time. Other facets relate to the different purposes for which MDS has been used, to various ways of looking at or "interpreting" an MDS representation, or to differences in the data required for the particular models. In this book, we give a fairly comprehensive presentation of MDS. For the reader with applied interests only, the first six chapters of Part I should be sufficient. They explain the basic notions of ordinary MDS, with an emphasis on how MDS can be helpful in answering substantive questions.

High-dimensional probability offers insight into the behavior of random vectors, random matrices, random subspaces, and objects used to quantify uncertainty in high dimensions. Drawing on ideas from probability, analysis, and geometry, it lends itself to applications in mathematics, statistics, theoretical computer science, signal processing, optimization, and more. It is the first to integrate theory, key tools, and modern applications of high-dimensional probability. Concentration inequalities form the core, and it covers both classical results such as Hoeffding's and Chernoff's inequalities and modern developments such as the matrix Bernstein's inequality. It then introduces the powerful methods based on stochastic processes, including such tools as Slepian's, Sudakov's, and Dudley's inequalities, as well as generic chaining and bounds based on VC dimension. A broad range of illustrations is embedded throughout, including classical and modern results for covariance estimation, clustering, networks, semidefinite programming, coding, dimension reduction, matrix completion, machine learning, compressed sensing, and sparse regression.

XML has become the lingua franca for representing business data, for exchanging information between business partners and applications, and for adding structure—and sometimes meaning—to text-based documents. XML offers some special challenges and opportunities in the area of search: querying XML can produce very precise, fine-grained results, if you know how to express and execute those queries. For software developers and systems architects: this book teaches the most useful approaches to querying XML documents and repositories. This book will also help managers and project leaders grasp how "querying XML fits into the larger context of querying and XML. Querying XML provides a comprehensive background from fundamental concepts (What is XML?) to data models (the Infoset, PSVI, XQuery Data Model), to APIs (querying XML from SQL or Java) and more. * Presents the concepts clearly, and demonstrates them with illustrations and examples; offers a thorough mastery of the subject area in a single book. * Provides comprehensive coverage of XML query languages, and the concepts needed to understand them completely (such as the XQuery Data Model). * Shows how to query XML documents and data using: XPath (the XML Path Language); XQuery, soon to be the new W3C Recommendation for querying XML; XQuery's companion XQueryX; and SQL, featuring the SQL/XML * Includes an extensive set of XQuery, XPath, SQL, Java, and other examples, with links to downloadable code and data samples.

This comprehensive and engaging textbook introduces the basic principles and techniques of signal processing, from the fundamental ideas of signals and systems theory to real-world applications. Students are introduced to the powerful foundations of modern signal processing, including the basic geometry of Hilbert space, the mathematics of Fourier transforms, and essentials of sampling, interpolation, approximation and compression. The authors discuss real-world issues and hurdles to using these tools, and ways of adapting them to overcome problems of finiteness and localization, the limitations of uncertainty, and computational costs. It includes over 160 homework problems and over 220 worked examples, specifically designed to test and expand students' understanding of the fundamentals of signal processing, and is accompanied by extensive online materials designed to aid learning, including Mathematica® resources and interactive demonstrations.

Software developers are faced with the challenge of making software systems and products of ever greater quality and safety, while at the same time being faced with the growing pressure of costs reduction in order to gain and maintain competitive advantages. As in any scientific and engineering discipline, reliable measurement is essential for talking on such a challenge. "Software measurement is an excellent abstraction mechanism for learning what works and what doesn't" (Victor Basili). Measurement of both software process and products provides a large amount of basic information for the evaluation of the software development processes or the software products themselves. Examples of recent successes in software measurement span multiple areas, such as evaluation of new development methods and paradigms, quality and management improvement programs, tool-supporting initiatives and company wide measurement programs. The German Computer Science Interest (GI) Group of Software Metrics and the Canadian Interest Group in Software Metrics (CIM) have attended to these concerns in the recent years. Research initiatives were directed initially to the definition of software metrics and then to validation of the software metrics themselves. This was followed by more and more investigation into practical applications of software metrics and by critical analysis of the benefits and weaknesses of software measurement programs. Key findings in this area of software engineering have been published in some important books, such as Dumke and Zuse's Theory and Practice of Software Measurement, Ebert and Dumke's Software Metrics in Practice and Lehner, Dumke and Abran's Software Metrics.

This treatise by an acknowledged expert includes several topics not found in any previous book.

The fundamental mathematical tools needed to understand machine learning include linear algebra, analytic geometry, matrix decompositions, vector calculus, optimization, probability and statistics. These topics are traditionally taught in disparate courses, making it

hard for data science or computer science students, or professionals, to efficiently learn the mathematics. This self-contained textbook bridges the gap between mathematical and machine learning texts, introducing the mathematical concepts with a minimum of prerequisites. It uses these concepts to derive four central machine learning methods: linear regression, principal component analysis, Gaussian mixture models and support vector machines. For students and others with a mathematical background, these derivations provide a starting point to machine learning texts. For those learning the mathematics for the first time, the methods help build intuition and practical experience with applying mathematical concepts. Every chapter includes worked examples and exercises to test understanding. Programming tutorials are offered on the book's web site.

This book is an English translation of the famous "Green Book" by Lafontaine and Pansu (1979). It has been enriched and expanded with new material to reflect recent progress. Additionally, four appendices, by Gromov on Levy's inequality, by Pansu on "quasiconvex" domains, by Katz on systoles of Riemannian manifolds, and by Semmes overviews analysis on metric spaces with measures, as well as an extensive bibliography and index round out this unique and beautiful book.

The digital compositing process is being applied in many diverse fields from Hollywood to corporate projects. Featuring over 30 pages of color, this tutorial/reference provides a complete overview of the technical and artistic skills necessary to undertake a digital composition project. The CD-ROM contains composition examples, illustrations, and development software.

The focus of these conference proceedings is on research, development, and applications in the fields of numerical geometry, scientific computing and numerical simulation, particularly in mesh generation and related problems. In addition, this year's special focus is on Voronoi diagrams and their applications, celebrating the 150th birthday of G.F. Voronoi. In terms of content, the book strikes a balance between engineering algorithms and mathematical foundations. It presents an overview of recent advances in numerical geometry, grid generation and adaptation in terms of mathematical foundations, algorithm and software development and applications. The specific topics covered include: quasi-conformal and quasi-isometric mappings, hyperelastic deformations, multidimensional generalisations of the equidistribution principle, discrete differential geometry, spatial and metric encodings, Voronoi-Delaunay theory for tilings and partitions, duality in mathematical programming and numerical geometry, mesh-based optimisation and optimal control methods. Further aspects examined include iterative solvers for variational problems and algorithm and software development. The applications of the methods discussed are multidisciplinary and include problems from mathematics, physics, biology, chemistry, material science, and engineering.

Until recently, almost all of the interactions between objects in virtual 3D worlds have been based on calculations performed using linear algebra. Linear algebra relies heavily on coordinates, however, which can make many geometric programming tasks very specific and complex-often a lot of effort is required to bring about even modest performance enhancements. Although linear algebra is an efficient way to specify low-level computations, it is not a suitable high-level language for geometric programming. Geometric Algebra for Computer Science presents a compelling alternative to the limitations of linear algebra. Geometric algebra, or GA, is a compact, time-effective, and performance-enhancing way to represent the geometry of 3D objects in computer programs. In this book you will find an introduction to GA that will give you a strong grasp of its relationship to linear algebra and its significance for your work. You will learn how to use GA to represent objects and perform geometric operations on them. And you will begin mastering proven techniques for making GA an integral part of your applications in a way that simplifies your code without slowing it down. * The first book on Geometric Algebra for programmers in computer graphics and entertainment computing * Written by leaders in the field providing essential information on this new technique for 3D graphics * This full colour book includes a website with GAViewer, a program to experiment with GA

Outlines a set of techniques that enable a researcher to discuss the "hidden structure" of large data bases. These techniques use proximities, measures which indicate how similar or different objects are, to find a configuration of points which reflects the structure in the data.

Several decades of psychometric research have led to the development of sophisticated models for multidimensional test data, and in recent years, multidimensional item response theory (MIRT) has become a burgeoning topic in psychological and educational measurement. Considered a cutting-edge statistical technique, the methodology underlying MIRT can be complex, and therefore doesn't receive much attention in introductory IRT courses. However author Wes Bonifay shows how MIRT can be understood and applied by anyone with a firm grounding in unidimensional IRT modeling. His volume includes practical examples and illustrations, along with numerous figures and diagrams. Multidimensional Item Response Theory includes snippets of R code interspersed throughout the text (with the complete R code included on an accompanying website) to guide readers in exploring MIRT models, estimating the model parameters, generating plots, and implementing the various procedures and applications discussed throughout the book.

"We finally have the definitive treatise on PyTorch! It covers the basics and abstractions in great detail. I hope this book becomes your extended reference document." —Soumith Chintala, co-creator of PyTorch Key Features Written by PyTorch's creator and key contributors Develop deep learning models in a familiar Pythonic way Use PyTorch to build an image classifier for cancer detection Diagnose problems with your neural network and improve training with data augmentation Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About The Book Every other day we hear about new ways to put deep learning to good use: improved medical imaging, accurate credit card fraud detection, long range weather forecasting, and more. PyTorch puts these superpowers in your hands. Instantly familiar to anyone who knows Python data tools like NumPy and Scikit-learn, PyTorch simplifies deep learning without sacrificing advanced features. It's great for building quick models, and it scales smoothly from laptop to enterprise. Deep Learning with PyTorch teaches you to create deep learning and neural network systems with PyTorch. This practical book gets you to work right away building a tumor image classifier from scratch. After covering the basics, you'll learn best practices for the entire deep learning pipeline, tackling advanced projects as your PyTorch skills become more sophisticated. All code samples are easy to explore in downloadable Jupyter notebooks. What You Will Learn Understanding deep learning data structures such as tensors and neural networks Best practices for the PyTorch Tensor API, loading data in Python, and visualizing results Implementing modules and loss functions Utilizing pretrained models from PyTorch Hub Methods for training networks with limited inputs Sifting through unreliable results to diagnose and fix problems in your neural network Improve your results with augmented data, better model architecture, and fine tuning This Book Is Written For For Python programmers with an interest in machine learning. No experience with PyTorch or other deep learning frameworks is required. About The Authors Eli Stevens has worked in Silicon Valley for the past 15 years as a software engineer, and the past 7 years as Chief Technical Officer of a startup making medical device software. Luca Antiga is co-founder and CEO of an AI engineering company located in Bergamo, Italy, and a regular contributor to PyTorch. Thomas Viehmann is a Machine Learning and PyTorch speciality trainer and consultant based in Munich, Germany and a PyTorch core developer. Table of Contents PART 1 - CORE PYTORCH 1 Introducing deep learning and the PyTorch Library 2 Pretrained networks 3 It starts with a tensor 4 Real-world data representation using tensors 5 The mechanics of learning 6 Using a neural network to fit the data 7 Telling birds from airplanes: Learning from images 8 Using convolutions to generalize PART 2 - LEARNING FROM IMAGES IN THE REAL WORLD: EARLY DETECTION OF LUNG CANCER 9 Using PyTorch to fight cancer 10 Combining data sources into a unified dataset 11 Training a classification model to detect suspected tumors 12 Improving training with metrics and augmentation 13 Using segmentation to find suspected nodules 14 End-to-end nodule analysis, and where to go next PART 3 - DEPLOYMENT 15 Deploying to production

A guide for constructing and using composite indicators for policy makers, academics, the media and other interested parties. In particular, this handbook is concerned with indicators which compare and rank country performance.

Each volume of Nicolas Bourbaki's well-known work, *The Elements of Mathematics*, contains a section or chapter devoted to the history of the subject. This book collects together those historical segments with an emphasis on the emergence, development, and interaction of the leading ideas of the mathematical theories presented in the *Elements*. In particular, the book provides a highly readable account of the evolution of algebra, geometry, infinitesimal calculus, and of the concepts of number and structure, from the Babylonian era through to the 20th century.

Foundations of Multidimensional and Metric Data Structures Morgan Kaufmann

The problem of privacy-preserving data analysis has a long history spanning multiple disciplines. As electronic data about individuals becomes increasingly detailed, and as technology enables ever more powerful collection and curation of these data, the need increases for a robust, meaningful, and mathematically rigorous definition of privacy, together with a computationally rich class of algorithms that satisfy this definition. Differential Privacy is such a definition. *The Algorithmic Foundations of Differential Privacy* starts out by motivating and discussing the meaning of differential privacy, and proceeds to explore the fundamental techniques for achieving differential privacy, and the application of these techniques in creative combinations, using the query-release problem as an ongoing example. A key point is that, by rethinking the computational goal, one can often obtain far better results than would be achieved by methodically replacing each step of a non-private computation with a differentially private implementation. Despite some powerful computational results, there are still fundamental limitations. Virtually all the algorithms discussed herein maintain differential privacy against adversaries of arbitrary computational power -- certain algorithms are computationally intensive, others are efficient. Computational complexity for the adversary and the algorithm are both discussed. The monograph then turns from fundamentals to applications other than query-release, discussing differentially private methods for mechanism design and machine learning. The vast majority of the literature on differentially private algorithms considers a single, static, database that is subject to many analyses. Differential privacy in other models, including distributed databases and computations on data streams, is discussed. *The Algorithmic Foundations of Differential Privacy* is meant as a thorough introduction to the problems and techniques of differential privacy, and is an invaluable reference for anyone with an interest in the topic.

Identifying, measuring and improving social impact is a significant challenge for corporate and private foundations, charities, NGOs and corporations. How best to balance possible social and environmental benefits (and costs) against one another? How does one bring clarity to multiple possibilities and opportunities? Based on years of work and new field studies from around the globe, the authors have written a book for managers that is grounded in the best academic and managerial research. It is a practical guide that describes the steps needed for identifying, measuring and improving social impact. This approach is useful in maximizing the impact of different types of investments, including grants and donations, impact investments, and commercial investments. With numerous examples of actual organizational approaches, research into more than fifty organizations, and extensive practical guidance and best practices, *Measuring and Improving Social Impacts* fills a critical gap.

This Fourth Edition introduces the latest theory and applications in optimization. It emphasizes constrained optimization, beginning with a substantial treatment of linear programming and then proceeding to convex analysis, network flows, integer programming, quadratic programming, and convex optimization. Readers will discover a host of practical business applications as well as non-business applications. Topics are clearly developed with many numerical examples worked out in detail. Specific examples and concrete algorithms precede more abstract topics. With its focus on solving practical problems, the book features free C programs to implement the major algorithms covered, including the two-phase simplex method, primal-dual simplex method, path-following interior-point method, and homogeneous self-dual methods. In addition, the author provides online JAVA applets that illustrate various pivot rules and variants of the simplex method, both for linear programming and for network flows. These C programs and JAVA tools can be found on the book's website. The website also includes new online instructional tools and exercises.

Poverty is a paradoxical state. Recognizable in the field for any sensitive observer who travels in remote rural areas and urban slums and meets marginalized people in a given society, poverty still remains a challenge to conceptual formalization and to measurement that is consistent with such formalization. The analysis of poverty is multidisciplinary. It goes from ethics to economics, from political science to human biology, and any type of measurement rests on mathematics. Moreover, poverty is multifaceted according to the types of deprivation, and it is also gender and age specific. A vector of variables is required, which raises a substantial problem for individual and group comparisons necessary to equity analysis. Multidimensionality also complicates the aggregation necessary to perform the efficiency analysis of policies. In the case of income poverty, these two problems, equity and efficiency, have benefited from very significant progress in the field of economics. Similar achievements are still to come in the area of multidimensional poverty. Within this general background, this book has a very modest and narrow-scoped objective. It proposes an operational methodology for measuring multidimensional poverty, independent from the conceptual origin, the size and the qualitative as well as the quantitative nature of the primary indicators used to describe the poverty of an individual, a household or a sociodemographic entity.

Publisher Description

Artificial Intelligence presents a practical guide to AI, including agents, machine learning and problem-solving simple and complex domains. This book provides an introduction to the mathematical and algorithmic foundations of data science, including machine learning, high-dimensional geometry, and analysis of large networks. Topics include the counterintuitive nature of data in high dimensions, important linear algebraic techniques such as singular value decomposition, the theory of random walks and Markov chains, the fundamentals of and important algorithms for machine learning, algorithms and analysis for clustering, probabilistic models for large networks, representation learning including topic modelling and non-negative matrix factorization, wavelets and compressed sensing. Important probabilistic techniques are developed including the law of large numbers, tail inequalities, analysis of random projections, generalization guarantees in machine learning, and moment methods for analysis of phase transitions in large random graphs. Additionally, important structural and complexity measures are discussed such as matrix norms and VC-dimension. This book is suitable for both undergraduate and graduate courses in the design and analysis of algorithms for data.

Balancing simplicity with technical rigour, this practical guide to the statistical techniques essential to research in marketing and related fields, describes each method as well as showing how they are applied. The book is accompanied by two real data sets to replicate examples and with exercises to solve, as well as detailed guidance on the use of appropriate software including: - 750 powerpoint slides with lecture notes and step-by-step guides to run analyses in SPSS (also includes screenshots) - 136 multiple choice questions for tests This is augmented by in-depth discussion of topics including: - Sampling - Data management and statistical packages - Hypothesis testing - Cluster analysis - Structural equation modelling

During the past decade there has been an explosion in computation and information technology. With it have come vast amounts of data in a variety of fields such as medicine, biology, finance, and marketing. The challenge of understanding these data has led to the development of new tools in the field of statistics, and spawned new areas such as data mining, machine learning, and bioinformatics. Many of these tools have common underpinnings but are often expressed with different terminology. This book describes the important ideas in these areas in a common conceptual framework. While the approach is statistical, the emphasis is on concepts rather than mathematics. Many examples are given, with a liberal use of color graphics. It should be a valuable resource for statisticians and anyone interested in data mining in science or industry. The book's coverage is broad, from supervised learning (prediction) to unsupervised learning. The many topics include neural

networks, support vector machines, classification trees and boosting--the first comprehensive treatment of this topic in any book. This major new edition features many topics not covered in the original, including graphical models, random forests, ensemble methods, least angle regression & path algorithms for the lasso, non-negative matrix factorization, and spectral clustering. There is also a chapter on methods for "wide" data (p bigger than n), including multiple testing and false discovery rates. Trevor Hastie, Robert Tibshirani, and Jerome Friedman are professors of statistics at Stanford University. They are prominent researchers in this area: Hastie and Tibshirani developed generalized additive models and wrote a popular book of that title. Hastie co-developed much of the statistical modeling software and environment in R/S-PLUS and invented principal curves and surfaces. Tibshirani proposed the lasso and is co-author of the very successful *An Introduction to the Bootstrap*. Friedman is the co-inventor of many data-mining tools including CART, MARS, projection pursuit and gradient boosting.

Winner of the 2006 Joseph W. Goodman Book Writing Award! A comprehensive treatment of the principles, mathematics, and statistics of image science In today's visually oriented society, images play an important role in conveying messages. From seismic imaging to satellite images to medical images, our modern society would be lost without images to enhance our understanding of our health, our culture, and our world. *Foundations of Image Science* presents a comprehensive treatment of the principles, mathematics, and statistics needed to understand and evaluate imaging systems. The book is the first to provide a thorough treatment of the continuous-to-discrete, or CD, model of digital imaging. *Foundations of Image Science* emphasizes the need for meaningful, objective assessment of image quality and presents the necessary tools for this purpose. Approaching the subject within a well-defined theoretical and physical context, this landmark text presents the mathematical underpinnings of image science at a level that is accessible to graduate students and practitioners working with imaging systems, as well as well-motivated undergraduate students. Destined to become a standard text in the field, *Foundations of Image Science* covers:

- Mathematical Foundations: Examines the essential mathematical foundations of image science
- Image Formation--Models and Mechanisms: Presents a comprehensive and unified treatment of the mathematical and statistical principles of imaging, with an emphasis on digital imaging systems and the use of SVD methods
- Image Quality: Provides a systematic exposition of the methodology for objective or task-based assessment of image quality
- Applications: Presents detailed case studies of specific direct and indirect imaging systems and provides examples of how to apply the various mathematical tools covered in the book
- Appendices: Covers the prerequisite material necessary for understanding the material in the main text, including matrix algebra, complex variables, and the basics of probability theory

Describes basic principles and recent developments in approximate query processing. It focuses on four key synopses: random samples, histograms, wavelets, and sketches. It considers issues such as accuracy, space and time efficiency, optimality, practicality, range of applicability, error bounds on query answers, and incremental maintenance.

This volume explores the scientific frontiers and leading edges of research across the fields of anthropology, economics, political science, psychology, sociology, history, business, education, geography, law, and psychiatry, as well as the newer, more specialized areas of artificial intelligence, child development, cognitive science, communications, demography, linguistics, and management and decision science. It includes recommendations concerning new resources, facilities, and programs that may be needed over the next several years to ensure rapid progress and provide a high level of returns to basic research.

[Copyright: d6804a9b99dd32d9332a76aed42a264b](#)