

Introduction To Data Analysis And Graphical Presentation In Biostatistics With R Statistics In The Large Springerbriefs In Statistics

Introduction to Functional Data Analysis provides a concise textbook introduction to the field. It explains how to analyze functional data, both at exploratory and inferential levels. It also provides a systematic and accessible exposition of the methodology and the required mathematical framework. The book can be used as textbook for a semester-long course on FDA for advanced undergraduate or MS statistics majors, as well as for MS and PhD students in other disciplines, including applied mathematics, environmental science, public health, medical research, geophysical sciences and economics. It can also be used for self-study and as a reference for researchers in those fields who wish to acquire solid understanding of FDA methodology and practical guidance for its implementation. Each chapter contains plentiful examples of relevant R code and theoretical and data analytic problems. The material of the book can be roughly divided into four parts of approximately equal length: 1) basic concepts and techniques of FDA, 2) functional regression models, 3) sparse and dependent functional data, and 4) introduction to the Hilbert space framework of FDA. The book assumes advanced undergraduate background in calculus, linear algebra, distributional probability theory, foundations of statistical inference, and some familiarity with R programming. Other required statistics background is provided in scalar settings before the related functional concepts are developed. Most chapters end with references to more advanced research for those who wish to gain a more in-depth understanding of a specific topic.

A practical guide to data-intensive humanities research using the Python programming language The use of quantitative methods in the humanities and related social sciences has increased considerably in recent years, allowing researchers to discover patterns in a vast range of source materials. Despite this growth, there are few resources addressed to students and scholars who wish to take advantage of these powerful tools. Humanities Data Analysis offers the first intermediate-level guide to quantitative data analysis for humanities students and scholars using the Python programming language. This practical textbook, which assumes a basic knowledge of Python, teaches readers the necessary skills for conducting humanities research in the rapidly developing digital environment. The book begins with an overview of the place of data science in the humanities, and proceeds to cover data carpentry: the essential techniques for gathering, cleaning, representing, and transforming textual and tabular data. Then, drawing from real-world, publicly available data sets that cover a variety of scholarly domains, the book delves into detailed case studies. Focusing on textual data analysis, the authors explore such diverse topics as network analysis, genre theory, onomastics, literacy, author attribution, mapping, stylometry, topic modeling, and time series analysis. Exercises and resources for further reading are provided at the end of each chapter. An ideal resource for humanities students and scholars aiming to take their Python skills to the next level, Humanities Data Analysis illustrates the benefits that quantitative methods can bring to complex research questions. Appropriate for advanced undergraduates, graduate students, and scholars with a basic knowledge of Python Applicable to many humanities disciplines, including history, literature, and sociology Offers real-world case studies using publicly available data sets Provides exercises at the end of each chapter for students to test acquired skills Emphasizes visual storytelling via data visualizations

This book covers several of the statistical concepts and data analytic skills needed to succeed in data-driven life science research. The authors proceed from relatively basic concepts related to computed p-values to advanced topics related to analyzing highthroughput data. They include the R code that performs this analysis and connect the lines of code to the statistical and mathematical concepts explained.

An introductory textbook offering a low barrier entry to data science; the hands-on approach will appeal to students from a range of disciplines.

The updated edition of this classic text introduces a range of techniques for exploring quantitative data. Beginning with an emphasis on descriptive statistics and graphical approaches, it moves on in later chapters to simple strategies for examining the associations between variables using inferential statistics such as chi squared. The book has been substantially revised to include the most recent approaches to data analysis, and includes step-by-step instructions on using SPSS. All these techniques are illustrated with intriguing real examples, drawn from important social research over the past three decades, designed to illuminate significant sociological and political debates. The book shows how students can use quantitative data to answer various questions: Is it true that the rich are getting richer and the poor are getting poorer? Are crime rates really going down, and how can we tell? How much alcohol do men and women really drink in an average week? Which country in Europe has the highest average working hours? Readers are encouraged to explore data for themselves, and are carefully guided through the opportunities and pitfalls of using statistical packages, as well as the numerous data sources readily available online. Suitable for those with no previous experience of quantitative data analysis, the second edition of Exploring Data will be invaluable to students across the social sciences. Visit the accompanying website at www.politybooks.com/exploringdata for more materials.

A guide to the principles and methods of data analysis that does not require knowledge of statistics or programming A General Introduction to Data Analytics is an essential guide to understand and use data analytics. This book is written using easy-to-understand terms and does not require familiarity with statistics or programming. The authors—noted experts in the field—highlight an explanation of the intuition behind the basic data analytics techniques. The text also contains exercises and illustrative examples. Thought to be easily accessible to non-experts, the book provides motivation to the necessity of analyzing data. It explains how to visualize and summarize data, and how to find natural groups and frequent patterns in a dataset. The book also explores predictive tasks, be them classification or regression. Finally, the book discusses popular data analytic applications,

like mining the web, information retrieval, social network analysis, working with text, and recommender systems. The learning resources offer: A guide to the reasoning behind data mining techniques A unique illustrative example that extends throughout all the chapters Exercises at the end of each chapter and larger projects at the end of each of the text's two main parts Together with these learning resources, the book can be used in a 13-week course guide, one chapter per course topic. The book was written in a format that allows the understanding of the main data analytics concepts by non-mathematicians, non-statisticians and non-computer scientists interested in getting an introduction to data science. A General Introduction to Data Analytics is a basic guide to data analytics written in highly accessible terms.

"This book introduces you to R, RStudio, and the tidyverse, a collection of R packages designed to work together to make data science fast, fluent, and fun. Suitable for readers with no previous programming experience"--

This introductory statistics textbook conveys the essential concepts and tools needed to develop and nurture statistical thinking. It presents descriptive, inductive and explorative statistical methods and guides the reader through the process of quantitative data analysis. In the experimental sciences and interdisciplinary research, data analysis has become an integral part of any scientific study. Issues such as judging the credibility of data, analyzing the data, evaluating the reliability of the obtained results and finally drawing the correct and appropriate conclusions from the results are vital. The text is primarily intended for undergraduate students in disciplines like business administration, the social sciences, medicine, politics, macroeconomics, etc. It features a wealth of examples, exercises and solutions with computer code in the statistical programming language R as well as supplementary material that will enable the reader to quickly adapt all methods to their own applications.

A multi-level introduction to Bayesian reasoning. The basic ideas of this approach to the quantification of uncertainty are presented using examples from research and everyday life. Applications covered include: parametric inference; combination of results; comparison of hypotheses; and more.

A valuable new edition of a standard reference The use of statistical methods for categorical data has increased dramatically, particularly for applications in the biomedical and social sciences. An Introduction to Categorical Data Analysis, Third Edition summarizes these methods and shows readers how to use them using software. Readers will find a unified generalized linear models approach that connects logistic regression and loglinear models for discrete data with normal regression for continuous data. Adding to the value in the new edition is: • Illustrations of the use of R software to perform all the analyses in the book • A new chapter on alternative methods for categorical data, including smoothing and regularization methods (such as the lasso), classification methods such as linear discriminant analysis and classification trees, and cluster analysis • New sections in many chapters introducing the Bayesian approach for the methods of that chapter • More than 70 analyses of data sets to illustrate application of the methods, and about 200 exercises, many containing other data sets • An appendix showing how to use SAS, Stata, and SPSS, and an appendix with short solutions to most odd-numbered exercises Written in an applied, nontechnical style, this book illustrates the methods using a wide variety of real data, including medical clinical trials, environmental questions, drug use by teenagers, horseshoe crab mating, basketball shooting, correlates of happiness, and much more. An Introduction to Categorical Data Analysis, Third Edition is an invaluable tool for statisticians and biostatisticians as well as methodologists in the social and behavioral sciences, medicine and public health, marketing, education, and the biological and agricultural sciences.

Introduction to Data Science: Data Analysis and Prediction Algorithms with R introduces concepts and skills that can help you tackle real-world data analysis challenges. It covers concepts from probability, statistical inference, linear regression, and machine learning. It also helps you develop skills such as R programming, data wrangling, data visualization, predictive algorithm building, file organization with UNIX/Linux shell, version control with Git and GitHub, and reproducible document preparation. This book is a textbook for a first course in data science. No previous knowledge of R is necessary, although some experience with programming may be helpful. The book is divided into six parts: R, data visualization, statistics with R, data wrangling, machine learning, and productivity tools. Each part has several chapters meant to be presented as one lecture. The author uses motivating case studies that realistically mimic a data scientist's experience. He starts by asking specific questions and answers these through data analysis so concepts are learned as a means to answering the questions. Examples of the case studies included are: US murder rates by state, self-reported student heights, trends in world health and economics, the impact of vaccines on infectious disease rates, the financial crisis of 2007-2008, election forecasting, building a baseball team, image processing of hand-written digits, and movie recommendation systems. The statistical concepts used to answer the case study questions are only briefly introduced, so complementing with a probability and statistics textbook is highly recommended for in-depth understanding of these concepts. If you read and understand the chapters and complete the exercises, you will be prepared to learn the more advanced concepts and skills needed to become an expert.

This accessible introduction to data analysis focuses on the interpretation of statistical results, in particular those which come from nonexperimental social research. It will provide social science researchers with the tools necessary to select and evaluate statistical tests appropriate for their research question. Using a consistent data-set throughout the book to illustrate the various analytic techniques, Michael Lewis-Beck covers topics such as univariate statistics, measures of association, the statistical significance of the relationship between two variables, simple regression in which the dependent variable is influenced by a single independent variable, and multiple regression.

Inverse problems are found in many applications, such as medical imaging, engineering, astronomy, and geophysics, among others. To solve an inverse problem is to recover an object from noisy, usually indirect observations. Solutions to inverse problems are subject to many potential sources of error introduced by approximate mathematical models, regularization methods, numerical approximations for efficient computations, noisy data, and limitations in the number of observations; thus it is important to include an assessment of the uncertainties as part of the solution. Such assessment is interdisciplinary by nature, as it requires, in addition to knowledge of the particular application, methods from applied mathematics, probability, and statistics. This book bridges applied mathematics and statistics by providing a basic introduction to probability and statistics for uncertainty quantification in the context of inverse problems, as well as an introduction to statistical regularization of inverse problems. The author covers basic statistical inference, introduces the framework of ill-posed inverse problems, and explains statistical questions that arise in their applications. An Introduction to Data Analysis and Uncertainty Quantification for Inverse Problems--includes many examples that explain techniques which are useful to address general problems arising in uncertainty quantification, Bayesian and non-Bayesian statistical methods and discussions of their complementary roles, and analysis of a real data set to illustrate the methodology covered throughout the book.

An Introduction to Data Science by Jeffrey S. Saltz and Jeffrey M. Stanton is an easy-to-read, gentle introduction for people with a wide range of backgrounds into the world of data science. Needing no prior coding experience or a deep understanding of statistics, this book uses the R programming language and RStudio® platform to make data science welcoming and accessible for all learners. After introducing

the basics of data science, the book builds on each previous concept to explain R programming from the ground up. Readers will learn essential skills in data science through demonstrations of how to use data to construct models, predict outcomes, and visualize data.

Through real-world datasets, this book shows the reader how to work with material in biostatistics using the open source software R. These include tools that are critical to dealing with missing data, which is a pressing scientific issue for those engaged in biostatistics. Readers will be equipped to run analyses and make graphical presentations based on the sample dataset and their own data. The hands-on approach will benefit students and ensure the accessibility of this book for readers with a basic understanding of R. Topics include: an introduction to Biostatistics and R, data exploration, descriptive statistics and measures of central tendency, t-Test for independent samples, t-Test for matched pairs, ANOVA, correlation and linear regression, and advice for future work.

The development of business intelligence has enhanced the visualization of data to inform and facilitate business management and strategizing. By implementing effective data-driven techniques, this allows for advance reporting tools to cater to company-specific issues and challenges. The Handbook of Research on Advanced Data Mining Techniques and Applications for Business Intelligence is a key resource on the latest advancements in business applications and the use of mining software solutions to achieve optimal decision-making and risk management results. Highlighting innovative studies on data warehousing, business activity monitoring, and text mining, this publication is an ideal reference source for research scholars, management faculty, and practitioners.

A guide for data managers and analyzers shares guidelines for identifying patterns, predicting future outcomes, and presenting findings to others; drawing on current research in cognitive science and learning theory while covering such additional topics as assessing data quality, handling ambiguous information, and organizing data within market groups. Original.

A fundamental book for social researchers. It provides a first-class, reliable guide to the basic issues in data analysis. Scholars and students can turn to it for teaching and applied needs with confidence.

Environmental Data Analysis with MatLab is for students and researchers working to analyze real data sets in the environmental sciences. One only has to consider the global warming debate to realize how critically important it is to be able to derive clear conclusions from often-noisy data drawn from a broad range of sources. This book teaches the basics of the underlying theory of data analysis, and then reinforces that knowledge with carefully chosen, realistic scenarios. MatLab, a commercial data processing environment, is used in these scenarios; significant content is devoted to teaching how it can be effectively used in an environmental data analysis setting. The book, though written in a self-contained way, is supplemented with data sets and MatLab scripts that can be used as a data analysis tutorial.

Author's website: <http://www.ideo.columbia.edu/users/menke/edawm/index.htm> Well written and outlines a clear learning path for researchers and students Uses real world environmental examples and case studies MatLab software for application in a readily-available software environment Homework problems help user follow up upon case studies with homework that expands them

A Hands-On Approach to Teaching Introductory Statistics Expanded with over 100 more pages, Introduction to Statistical Data Analysis for the Life Sciences, Second Edition presents the right balance of data examples, statistical theory, and computing to teach introductory statistics to students in the life sciences. This popular textbook covers the m

Covering the general process of data analysis to finding, collecting, organizing, and presenting data, this book offers a complete introduction to the fundamentals of data analysis. Using real-world case studies as illustrations, it helps readers understand theories behind and develop techniques for conducting quantitative, qualitative, and mixed methods data analysis. With an easy-to-follow organization and clear, jargon-free language, it helps readers not only become proficient data analysts, but also develop the critical thinking skills necessary to assess analyses presented by others in both academic research and the popular media. It includes advice on: - Data analysis frameworks - Validity and credibility of data - Sampling techniques - Data management - The big data phenomenon - Data visualisation - Effective data communication Whether you are new to data analysis or looking for a quick-reference guide to key principles of the process, this book will help you uncover nuances, complexities, patterns, and relationships among all types of data.

An introduction to statistical data mining, Data Analysis and Data Mining is both textbook and professional resource. Assuming only a basic knowledge of statistical reasoning, it presents core concepts in data mining and exploratory statistical models to students and professional statisticians-both those working in communications and those working in a technological or scientific capacity-who have a limited knowledge of data mining. This book presents key statistical concepts by way of case studies, giving readers the benefit of learning from real problems and real data. Aided by a diverse range of statistical methods and techniques, readers will move from simple problems to complex problems. Through these case studies, authors Adelchi Azzalini and Bruno Scarpa explain exactly how statistical methods work; rather than relying on the "push the button" philosophy, they demonstrate how to use statistical tools to find the best solution to any given problem. Case studies feature current topics highly relevant to data mining, such web page traffic; the segmentation of customers; selection of customers for direct mail commercial campaigns; fraud detection; and measurements of customer satisfaction. Appropriate for both advanced undergraduate and graduate students, this much-needed book will fill a gap between higher level books, which emphasize technical explanations, and lower level books, which assume no prior knowledge and do not explain the methodology behind the statistical operations.

Ott and Longnecker's AN INTRODUCTION TO STATISTICAL METHODS AND DATA ANALYSIS, 6th Edition, International Edition provides a broad overview of statistical methods for advanced undergraduate and graduate students from a variety of disciplines who have little or no prior course work in statistics. The authors teach students to solve problems encountered in research projects, to make decisions based on data in general settings both within and beyond the university setting, and to become critical readers of statistical analyses in research papers and in news reports. The first eleven chapters present material typically covered in an introductory statistics course, as well as case studies and examples that are often encountered in undergraduate capstone courses. The remaining chapters cover regression modeling and design of experiments.

DATA ANALYTICS FOR BEGINNER: IN ORDER TO SUCEED IN TODAY'Ss FAST PACE BUSINESS ENVIRONEMNT, YOU NEED TO MASTER DATA ANALYTICS. Data Analytics is the most powerful tool to analyze today's business environment and to predict future developments. Is it not the dream of every business owner to know exactly what the customer will buy in 6 months or what the new product hype will look like in your OWN industry? Data Analytics is the tool that will bring you answers to these questions. Here's why Data Analytics for Beginners will bring your business to a complete new level:

How you can use data analytics to improve your business How to plan data analysis to know exactly what your target group wants How to implement descriptive analysis You will learn the exact techniques that are required to master Data Analytics Our customer's feedback I am the owner of a home supplies shop with 15 employees and this book improved the sales by 18,5% during the last 3 months. Richard S., Boston. Data Analytics for Beginners was a eye opener for me and my business. With this book I research all of my products on sale and my skills about the market I am in enhanced drastically. I can recommend this book to everyone that is planning to improve the business. Anamda R., Sacramento. During my IT studies this book supported me a lot with anaylsis about future business trends. This book has a easy to understand writing style without any expert language. In other words: every beginner can work with this book right away. Thomas E., Baltimore. Here's what you will get Planning a Study Surveys Experiments Gathering Data How to select useful samples Avoiding Bias in Data Sets Descriptive Analysis Mean Median Mode Variance Standard Deviation Coefficient of Variation Pie Charts How to create Pie Charts in Excel Bar Graphs How to Create Bar Charts in Excel Time Charts and Line Charts How to create a time chart in excel How to create a line chart in excel Histograms How to create a

histogram in Excel Scatter Plots How to create a Scatter Chart in Excel Business Intelligence Data Analytics in Business and Industry

Introduction to Data Science: Data Analysis and Prediction Algorithms with R introduces concepts and skills that can help you tackle real-world data analysis challenges. It covers concepts from probability, statistical inference, linear regression, and machine learning. It also helps you develop skills such as R programming, data wrangling, data visualization, predictive algorithm building, file organization with UNIX/Linux shell, version control with Git and GitHub, and reproducible document preparation. This book is a textbook for a first course in data science. No previous knowledge of R is necessary, although some experience with programming may be helpful. The book is divided into six parts: R, data visualization, statistics with R, data wrangling, machine learning, and productivity tools. Each part has several chapters meant to be presented as one lecture. The author uses motivating case studies that realistically mimic a data scientist's experience. He starts by asking specific questions and answers these through data analysis so concepts are learned as a means to answering the questions. Examples of the case studies included are: US murder rates by state, self-reported student heights, trends in world health and economics, the impact of vaccines on infectious disease rates, the financial crisis of 2007-2008, election forecasting, building a baseball team, image processing of hand-written digits, and movie recommendation systems. The statistical concepts used to answer the case study questions are only briefly introduced, so complementing with a probability and statistics textbook is highly recommended for in-depth understanding of these concepts. If you read and understand the chapters and complete the exercises, you will be prepared to learn the more advanced concepts and skills needed to become an expert. A complete solutions manual is available to registered instructors who require the text for a course.

Statistical methods provide a logical, coherent framework in which data from experimental science can be analyzed. However, many researchers lack the statistical skills or resources that would allow them to explore their data to its full potential. Introduction to Data Analysis with R for Forensic Sciences minimizes theory and mathematics and focuses on the application and practice of statistics to provide researchers with the dexterity necessary to systematically analyze data discovered from the fruits of their research. Using traditional techniques and employing examples and tutorials with real data collected from experiments, this book presents the following critical information necessary for researchers: A refresher on basic statistics and an introduction to R Considerations and techniques for the visual display of data through graphics An overview of statistical hypothesis tests and the reasoning behind them A comprehensive guide to the use of the linear model, the foundation of most statistics encountered An introduction to extensions to the linear model for commonly encountered scenarios, including logistic and Poisson regression Instruction on how to plan and design experiments in a way that minimizes cost and maximizes the chances of finding differences that may exist Focusing on forensic examples but useful for anyone working in a laboratory, this volume enables researchers to get the most out of their experiments by allowing them to cogently analyze the data they have collected, saving valuable time and effort.

Data Mining and Analytics provides a broad and interactive overview of a rapidly growing field. The exponentially increasing rate at which data is generated creates a corresponding need for professionals who can effectively handle its storage, analysis, and translation.

This accessible and classroom-tested textbook/reference presents an introduction to the fundamentals of the emerging and interdisciplinary field of data science. The coverage spans key concepts adopted from statistics and machine learning, useful techniques for graph analysis and parallel programming, and the practical application of data science for such tasks as building recommender systems or performing sentiment analysis. Topics and features: provides numerous practical case studies using real-world data throughout the book; supports understanding through hands-on experience of solving data science problems using Python; describes techniques and tools for statistical analysis, machine learning, graph analysis, and parallel programming; reviews a range of applications of data science, including recommender systems and sentiment analysis of text data; provides supplementary code resources and data at an associated website.

Data Science in Education Using R is the go-to reference for learning data science in the education field. The book answers questions like: What does a data scientist in education do? How do I get started learning R, the popular open-source statistical programming language? And what does a data analysis project in education look like? If you're just getting started with R in an education job, this is the book you'll want with you. This book gets you started with R by teaching the building blocks of programming that you'll use many times in your career. The book takes a "learn by doing" approach and offers eight analysis walkthroughs that show you a data analysis from start to finish, complete with code for you to practice with. The book finishes with how to get involved in the data science community and how to integrate data science in your education job. This book will be an essential resource for education professionals and researchers looking to increase their data analysis skills as part of their professional and academic development.

Modern Data Analysis contains the proceedings of a Workshop on Modern Data Analysis held in Raleigh, North Carolina, on June 2-4, 1980 under the auspices of the United States Army Research Office. The papers review theories and methods of data analysis and cover topics ranging from single and multiple quantile-quantile (Q-Q) plotting procedures to biplot display and pencil-and-paper exploratory data analysis methods. Projection pursuit methods for data analysis are also discussed. Comprised of nine chapters, this book begins with an introduction to styles of data analysis techniques, followed by an analysis of single and multiple Q-Q plotting procedures. Problems involving extreme-value data and the behavior of sample averages are considered. Subsequent chapters deal with the use of smelting in guiding re-expression; geometric data analysis; and influence functions and regression diagnostics. The final chapter examines the use and interpretation of robust analysis of variance for the general non-full-rank linear model. The procedures are described in terms of their mathematical structure, which leads to efficient computational algorithms. This monograph should be of interest to mathematicians and statisticians.

An Introduction to Statistics and Data Analysis Using Stata® by Lisa Daniels and Nicholas Minot provides a step-by-step introduction for statistics, data analysis, or research methods classes with Stata. Concise descriptions emphasize the concepts behind statistics for students rather than the derivations of the formulas. With real-world examples from a variety of disciplines and extensive detail on the commands in Stata, this text provides an integrated approach to research design, statistical analysis, and report writing for social science students.

Data Analytics: A Small Data Approach is suitable for an introductory data analytics course to help students understand some main statistical learning models. It has many small datasets to guide students to work out pencil solutions of the models and then compare with results obtained from established R packages. Also, as data science practice is a process that should be told as a story, in this book there are many course materials about exploratory data analysis, residual analysis, and flowcharts to develop and validate models and data pipelines. The main models covered in this book include linear regression, logistic regression, tree models and random forests, ensemble learning, sparse learning, principal component analysis, kernel methods including the support vector machine and kernel regression, and deep learning. Each chapter introduces two or three techniques. For each technique, the book highlights the intuition and rationale first, then shows how mathematics is used to articulate the intuition and formulate the learning problem. R is used to implement the techniques on both simulated and real-world dataset. Python code is also available at the book's website: <http://dataanalyticsbook.info>.

Get complete instructions for manipulating, processing, cleaning, and crunching datasets in Python. Updated for Python 3.6, the second edition of this hands-on guide is packed with practical case studies that show you how to solve a broad set of data analysis problems effectively. You'll learn the latest versions of pandas, NumPy, IPython, and Jupyter

in the process. Written by Wes McKinney, the creator of the Python pandas project, this book is a practical, modern introduction to data science tools in Python. It's ideal for analysts new to Python and for Python programmers new to data science and scientific computing. Data files and related material are available on GitHub. Use the IPython shell and Jupyter notebook for exploratory computing Learn basic and advanced features in NumPy (Numerical Python) Get started with data analysis tools in the pandas library Use flexible tools to load, clean, transform, merge, and reshape data Create informative visualizations with matplotlib Apply the pandas groupby facility to slice, dice, and summarize datasets Analyze and manipulate regular and irregular time series data Learn how to solve real-world data analysis problems with thorough, detailed examples

This book is a comprehensive introduction to the methods and algorithms of modern data analytics. It provides a sound mathematical basis, discusses advantages and drawbacks of different approaches, and enables the reader to design and implement data analytics solutions for real-world applications. This book has been used for more than ten years in the Data Mining course at the Technical University of Munich. Much of the content is based on the results of industrial research and development projects at Siemens.

This textbook offers an easy-to-follow, practical guide to modern data analysis using the programming language R. The chapters cover topics such as the fundamentals of programming in R, data collection and preprocessing, including web scraping, data visualization, and statistical methods, including multivariate analysis, and feature exercises at the end of each section. The text requires only basic statistics skills, as it strikes a balance between statistical and mathematical understanding and implementation in R, with a special emphasis on reproducible examples and real-world applications. This textbook is primarily intended for undergraduate students of mathematics, statistics, physics, economics, finance and business who are pursuing a career in data analytics. It will be equally valuable for master students of data science and industry professionals who want to conduct data analyses.

Discover how data science can help you gain in-depth insight into your business - the easy way! Jobs in data science abound, but few people have the data science skills needed to fill these increasingly important roles. Data Science For Dummies is the perfect starting point for IT professionals and students who want a quick primer on all areas of the expansive data science space. With a focus on business cases, the book explores topics in big data, data science, and data engineering, and how these three areas are combined to produce tremendous value. If you want to pick-up the skills you need to begin a new career or initiate a new project, reading this book will help you understand what technologies, programming languages, and mathematical methods on which to focus. While this book serves as a wildly fantastic guide through the broad, sometimes intimidating field of big data and data science, it is not an instruction manual for hands-on implementation. Here's what to expect: Provides a background in big data and data engineering before moving on to data science and how it's applied to generate value Includes coverage of big data frameworks like Hadoop, MapReduce, Spark, MPP platforms, and NoSQL Explains machine learning and many of its algorithms as well as artificial intelligence and the evolution of the Internet of Things Details data visualization techniques that can be used to showcase, summarize, and communicate the data insights you generate It's a big, big data world out there—let Data Science For Dummies help you harness its power and gain a competitive edge for your organization.

[Copyright: 1ce7b1a00c3936cea8d0ea63d6b63108](https://www.springer.com/9781493993666)