Big Data Does Size Matter

Big Data in Radio Astronomy: Scientific Data Processing for Advanced Radio Telescopes provides the latest research developments in big data methods and techniques for radio astronomy. Providing examples from such projects as the Square Kilometer Array (SKA), the world's largest radio telescope that generates over an Exabyte of data every day, the book offers solutions for coping with the challenges and opportunities presented by the exponential growth of astronomical data. Presenting state-of-the-art results and research, this book is a timely reference for both practitioners and researchers working in radio astronomy, as well as students looking for a basic understanding of big data in astronomy. Bridges the gap between radio astronomy and computer science Includes coverage of the observation lifecycle as well as data collection, processing and analysis Presents state-of-the-art research and techniques in big data related to radio astronomy Utilizes real-world examples, such as Square Kilometer Array (SKA) and Five-hundred-meter Aperture Spherical radio Telescope (FAST)

A concise introduction to the emerging field of data science, explaining its evolution, relation to machine learning, current uses, data infrastructure issues, and ethical challenges. The goal of data science is to improve decision making through the analysis of data. Today data science determines the ads we see online, the books and movies that are recommended to us online, which emails are filtered into our spam folders, and even how much we pay for health insurance. This volume in the MIT Press Essential Knowledge series offers a concise introduction to the emerging field of data science, explaining its evolution, current uses, data infrastructure issues, and ethical challenges. It has never been easier for organizations to gather, store, and process data. Use of data science is driven by the rise of big data and social media, the development of high-performance computing, and the emergence of such powerful methods for data analysis and modeling as deep learning. Data science encompasses a set of principles, problem definitions, algorithms, and processes for extracting non-obvious and useful patterns from large datasets. It is closely related to the fields of data mining and machine learning, but broader in scope. This book offers a brief history of the field, introduces fundamental data concepts, and describes the stages in a data science project. It considers data infrastructure and the challenges posed by integrating data from multiple sources, introduces the basics of machine learning, and discusses how to link machine learning expertise with real-world problems. The book also reviews ethical and legal issues, developments in data regulation, and computational approaches to preserving privacy. Finally, it considers the future impact of data science and offers principles for success in data science projects.

The world is witnessing the growth of a global movement facilitated by technology and social media. Fueled by information, this movement contains enormous potential to create more accountable, efficient, responsive, and effective governments and businesses, as well as spurring economic growth. Big Data Governance and Perspectives in Knowledge Management is a collection of innovative research on the methods and applications of applying robust processes around data, and aligning organizations and skillsets around those processes. Highlighting a range of topics including data analytics, prediction analysis, and software development, this book is ideally designed for academicians, researchers, information science professionals, software developers, computer engineers, graduate-level computer science students, policymakers, and managers seeking current research on the convergence of big data and information governance as two major trends in information management.

A revelatory exploration of the hottest trend in technology and the dramatic impact it will have on the economy, science, and society at large. Which paint color is most likely to tell you that a used car is in good shape? How can officials identify the most dangerous New York City manholes before they explode? And how did Google searches predict the spread of the H1N1 flu outbreak? The key to answering these questions, and many more, is big data. "Big data" refers to our burgeoning ability to crunch vast collections of information, analyze it instantly, and draw sometimes profoundly surprising conclusions from it. This emerging science can translate myriad phenomena—from the price of airline tickets to the text of millions of books—into searchable form, and uses our increasing computing power to unearth epiphanies that we never could have seen before. A revolution on par with the Internet or perhaps even the printing press, big data will change the way we think about business, health, politics, education, and innovation in the years to come. It also poses fresh threats, from the inevitable end of privacy as we know it to the prospect of being penalized for things we haven't even done yet, based on big data's ability to predict our future behavior. In this brilliantly clear, often surprising work, two leading experts explain what big data is, how it will change our lives, and what we can do to protect ourselves from its hazards.

Big Data is the first big book about the next big thing. www.big-data-book.com

As technology advances, high volumes of valuable data are generated day by day in modern organizations. The management of such huge volumes of data has become a priority in these organizations, requiring new techniques for data management and data analysis in Big Data environments. These environments encompass many different fields including medicine, education data, and recommender systems. The aim of this book is to provide the reader with a variety of fields and systems where the analysis and management of Big Data are essential. This book describes the importance of the Big Data era and how existing information systems are required to be adapted to face up the problems derived from the management of massive datasets.

Big Data represents a new era in data exploration and utilization, and IBM is uniquely positioned to help clients navigate this transformation. This book reveals how IBM is leveraging open source Big Data technology, infused with IBM technologies, to deliver a robust, secure, highly available, enterprise-class Big Data platform. The three defining characteristics of Big Data--volume, variety, and velocity--are discussed. You'll get a primer on Hadoop and how IBM is hardening it for the enterprise, and learn when to leverage IBM InfoSphere BigInsights (Big Data at rest) and IBM InfoSphere Streams (Big Data in motion) technologies. Industry use cases are also included in this practical guide. Learn how IBM hardens Hadoop for enterprise-class scalability and reliability Gain insight into IBM's unique in-motion and at-rest Big Data analytics platform Learn tips and tricks for Big Data use cases and solutions Get a quick Hadoop primer

Is the Brexit vote successful big data politics or the end of democracy? Why do airlines overbook, and why do banks get it wrong so often? How does big data enable Netflix to forecast a hit, CERN to find the Higgs boson and medics to discover if red wine really is good for you? And how are companies using big data to benefit from smart meters, use advertising that spies on you and develop the gig economy, where workers are managed by the whim of an algorithm? The volumes of data we now access can give unparalleled abilities to make predictions, respond to customer demand and solve problems. But Big Brother's shadow hovers over it. Though big data can set us free and enhance our lives, it has the potential to create an underclass and a totalitarian state. With big data ever-present, you can't afford to ignore it. Acclaimed science writer Brian Clegg - a habitual early adopter of new technology (and the owner of the second-ever copy of Windows in the UK) - brings big data to life.

Thinking Big Data in Geography offers a practical state-of-the-field overview of big data as both a means and an object of research, with essays from prominent and emerging scholars such as Rob Kitchin, Renee Sieber, and Mark Graham. Part 1 explores how the advent of geoweb technologies and big data sets has influenced some of geography's major subdisciplines: urban politics and political economy, human-environment interactions, and geographic information sciences. Part 2 addresses how the geographic study of big data has implications for other disciplinary fields, notably the digital humanities and the study of social justice. The volume concludes with theoretical applications of the geoweb and big data as they pertain to society as a whole, examining the ways in which user-generated data come into the world and are complicit in its unfolding. The contributors raise caution regarding the use of spatial big data, citing issues of accuracy, surveillance, and privacy.

A handy reference guide for data analysts and data scientists to help to obtain value from big data analytics using Spark on Hadoop clusters About This Book This book is based

on the latest 2.0 version of Apache Spark and 2.7 version of Hadoop integrated with most commonly used tools. Learn all Spark stack components including latest topics such as DataFrames, DataSets, GraphFrames, Structured Streaming, DataFrame based ML Pipelines and SparkR. Integrations with frameworks such as HDFS, YARN and tools such as Jupyter, Zeppelin, NiFi, Mahout, HBase Spark Connector, GraphFrames, H2O and Hivemall. Who This Book Is For Though this book is primarily aimed at data analysts and data scientists, it will also help architects, programmers, and practitioners. Knowledge of either Spark or Hadoop would be beneficial. It is assumed that you have basic programming background in Scala, Python, SQL, or R programming with basic Linux experience. Working experience within big data environments is not mandatory. What You Will Learn Find out and implement the tools and techniques of big data analytics using Spark on Hadoop clusters with wide variety of tools used with Spark and Hadoop Understand all the Hadoop and Spark ecosystem components Get to know all the Spark components: Spark Core, Spark SQL, DataFrames, DataSets, Conventional and Structured Streaming, MLLib, ML Pipelines and Graphx See batch and real-time data analytics using Spark Core, Spark SQL, and Conventional and Structured Streaming Get to grips with data science and machine learning using MLLib, ML Pipelines, H2O, Hivemall, Graphx, SparkR and Hivemall. In Detail Big Data Analytics book aims at providing the fundamentals of Apache Spark and Hadoop. All Spark components – Spark Core, Spark SQL, DataFrames, Data sets, Conventional Streaming, Structured Streaming, MLlib, Graphx and Hadoop core components – HDFS, MapReduce and Yarn are explored in greater depth with implementation examples on Spark + Hadoop clusters. It is moving away from MapReduce to Spark. So, advantages of Spark over MapReduce are explained at great depth to reap benefits of in-memory speeds. DataFrames API, Data Sources API and new Data set API are explained for building Big Data analytical applications. Real-time data analytics using Spark Streaming with Apache Kafka and HBase is covered to help building streaming applications. New Structured streaming concept is explained with an IOT (Internet of Things) use case. Machine learning techniques are covered using MLLib, ML Pipelines and SparkR and Graph Analytics are covered with GraphX and GraphFrames components of Spark. Readers will also get an opportunity to get started with web based notebooks such as Jupyter, Apache Zeppelin and data flow tool Apache NiFi to analyze and visualize data. Style and approach This step-by-step pragmatic guide will make life easy no matter what your level of experience. You will deep dive into Apache Spark on Hadoop clusters through ample exciting real-life examples. Practical tutorial explains data science in simple terms to help programmers and data analysts get started with Data Science

Longlisted for the National Book Award New York Times Bestseller A former Wall Street quant sounds an alarm on the mathematical models that pervade modern life -- and threaten to rip apart our social fabric We live in the age of the algorithm. Increasingly, the decisions that affect our lives--where we go to school, whether we get a car loan, how much we pay for health insurance--are being made not by humans, but by mathematical models. In theory, this should lead to greater fairness: Everyone is judged according to the same rules, and bias is eliminated. But as Cathy O'Neil reveals in this urgent and necessary book, the opposite is true. The models being used today are opaque, unregulated, and uncontestable, even when they're wrong. Most troubling, they reinforce discrimination: If a poor student can't get a loan because a lending model deems him too risky (by virtue of his zip code), he's then cut off from the kind of education that could pull him out of poverty, and a vicious spiral ensues. Models are propping up the lucky and punishing the downtrodden, creating a "toxic cocktail for democracy." Welcome to the dark side of Big Data. Tracing the arc of a person's life, O'Neil exposes the black box models that shape our future, both as individuals and as a society. These "weapons of math destruction" score teachers and students, sort r sum s, grant (or deny) loans, evaluate workers, target voters, set parole, and monitor our health. O'Neil calls on modelers to take more responsibility for their algorithms and on policy makers to regulate their use. But in the end, it's up to us to become more savvy about the models that govern our lives. This important book empowers us to ask the tough questions, uncover the truth, and demand change. -- Longlist for National Book Award (Non-Fiction) -- Goodreads, semi-finalist for the 2016 Goodreads Choice Awards (Science and Technology) -- Kirkus, Best Books of 2016 -- New York Times, 100 Notable Books of 2016 (Non-Fiction) -- The Guardian, Best Books of 2016 -- WBUR's "On Point," Best Books of 2016: Staff Picks -- Boston Globe, Best Books of 2016, Non-Fiction

Go ahead, be skeptical about big data. The author was—at first. When the term “big data” first came on the scene, bestselling author Tom Davenport (Competing on Analytics, Analytics at Work) thought it was just another example of technology hype. But his research in the years that followed changed his mind. Now, in clear, conversational language, Davenport explains what big data means—and why everyone in business needs to know about it. Big Data at Work covers all the bases: what big data means from a technical, consumer, and management perspective; what its opportunities and costs are; where it can have real business impact; and which aspects of this hot topic have been oversold. This book will help you understand: • Why big data is important to you and your organization • What technology you need to manage it • How big data could change your job, your company, and your industry • How to hire, rent, or develop the kinds of people who make big data work • The key success factors in implementing any big data project • How big data is leading to a new approach to managing analytics With dozens of company examples, including UPS, GE, Amazon, United Healthcare, Citigroup, and many others, this book will help you seize all opportunities—from improving decisions, products, and services to strengthening customer relationships. It will show you how to put big data to work in your own organization so that you too can harness the power of this ever-evolving new resource.

Big Data: A Business and Legal Guide supplies a clear understanding of the interrelationships between Big Data, the new business insights it reveals, and the laws, regulations, and contracting practices that impact the use of the insights and the data. Providing business executives and lawyers (in-house and in private practice) with an accessible primer on Big Data and its business implications, this book will enable readers to quickly grasp the key issues and effectively implement the right solutions to collecting, licensing,

handling, and using Big Data. The book brings together subject matter experts who examine a different area of law in each chapter and explain how these laws can affect the way your business or organization can use Big Data. These experts also supply recommendations as to the steps your organization can take to maximize Big Data opportunities without increasing risk and liability to your organization. Provides a new way of thinking about Big Data that will help readers address emerging issues Supplies real-world advice and practical ways to handle the issues Uses examples pulled from the news and cases to illustrate points Includes a non-technical Big Data primer that discusses the characteristics of Big Data and distinguishes it from traditional database models Taking a cross-disciplinary approach, the book will help executives, managers, and counsel better understand the interrelationships between Big Data, decisions based on Big Data, and the laws, regulations, and contracting practices that impact its use. After reading this book, you will be able to think more broadly about the best way to harness Big Data in your business and establish procedures to ensure that legal considerations are part of the decision.

A new way of thinking about data science and data ethics that is informed by the ideas of intersectional feminism. Today, data science is a form of power. It has been used to expose injustice, improve health outcomes, and topple governments. But it has also been used to discriminate, police, and surveil. This potential for good, on the one hand, and harm, on the other, makes it essential to ask: Data science by whom? Data science for whom? Data science with whose interests in mind? The narratives around big data and data science are overwhelmingly white, male, and techno-heroic. In Data Feminism, Catherine D'Ignazio and Lauren Klein present a new way of thinking about data science and data ethics—one that is informed by intersectional feminist thought. Illustrating data feminism in action, D'Ignazio and Klein show how challenges to the male/female binary can help challenge other hierarchical (and empirically wrong) classification systems. They explain how, for example, an understanding of emotion can expand our ideas about effective data visualization, and how the concept of invisible labor can expose the significant human efforts required by our automated systems. And they show why the data never, ever “speak for themselves.” Data Feminism offers strategies for data scientists seeking to learn how feminism can help them work toward justice, and for feminists who want to focus their efforts on the growing field of data science. But Data Feminism is about much more than gender. It is about power, about who has it and who doesn't, and about how those differentials of power can be challenged and changed.

The Structure of Digital Computing takes a fifty year perspective on computing and discusses what is significant, what is novel, what endures, and why it is all so confusing. The book tries to balance two point of views: digital computing as viewed from a business perspective, where the focus is on marketing and selling, and digital computing from a research perspective, where the focus is on developing fundamentally new technology.

Introduces professionals and scientists to statistics and machine learning using the programming language R Written by and for practitioners, this book provides an overall introduction to R, focusing on tools and methods commonly used in data science, and placing emphasis on practice and business use. It covers a wide range of topics in a single volume, including big data, databases, statistical machine learning, data wrangling, data visualization, and the reporting of results. The topics covered are all important for someone with a science/math background that is looking to quickly learn several practical technologies to enter or transition to the growing field of data science. The Big R-Book for Professionals: From Data Science to Learning Machines and Reporting with R includes nine parts, starting with an introduction to the subject and followed by an overview of R and elements of statistics. The third part revolves around data, while the fourth focuses on data wrangling. Part 5 teaches readers about exploring data. In Part 6 we learn to build models, Part 7 introduces the reader to the reality in companies, Part 8 covers reports and interactive applications and finally Part 9 introduces the reader to big data and performance computing. It also includes some helpful appendices. Provides a practical guide for non-experts with a focus on business users Contains a unique combination of topics including an introduction to R, machine learning, mathematical models, data wrangling, and reporting Uses a practical tone and integrates multiple topics in a coherent framework Demystifies the hype around machine learning and AI by enabling readers to understand the provided models and program them in R Shows readers how to visualize results in static and interactive reports Supplementary materials includes PDF slides based on the book’s content, as well as all the extracted R-code and is available to everyone on a Wiley Book Companion Site The Big R-Book is an excellent guide for science technology, engineering, or mathematics students who wish to make a successful transition from the academic world to the professional. It will also appeal to all young data scientists, quantitative analysts, and analytics professionals, as well as those who make mathematical models.

Big Data Mining for Climate Change addresses how to manage the vast amount of information available for analysis. Climate change and its environmental, economic and social consequences are widely recognized as the biggest, most interconnected problem facing humanity. There is a huge amount of potential information currently available...and it is growing exponentially. This book walks through the latest research and how to navigate the resources available using big data applications. It is appropriate for scientists and advanced students studying climate change from a number of disciplines, including the atmospheric sciences, oceanic sciences, geography, environment sciences, ecology, energy, economics, engineering and public policy. Provides a step-by-step guide for applying big data mining tools to climate and environmental research Presents a comprehensive review of theory and algorithms of big data mining for climate change Includes current research in climate and environmental science as it relates to using big data algorithms

What is Big Data, and why should you care? Big data knows where you've been and who your friends are. It knows what you like and what makes you angry. It can predict what

you'll buy, where you'll be the victim of crime and when you'll have a heart attack. Big data knows you better than you know yourself, or so it claims. But how well do you know big data? You've probably seen the phrase in newspaper headlines, at work in a marketing meeting, or on a fitness-tracking gadget. But can you understand it without being a Silicon Valley nerd who writes computer programs for fun? Yes. Yes, you can. Timandra Harkness writes comedy, not computer code. The only programmes she makes are on the radio. If you can read a newspaper you can read this book. Starting with the basics – what IS data? And what makes it big? – Timandra takes you on a whirlwind tour of how people are using big data today: from science to smart cities, business to politics, self-quantification to the Internet of Things. Finally, she asks the big questions about where it's taking us; is it too big for its boots, or does it think too small? Are you a data point or a human being? Will this book be full of rhetorical questions? No. It also contains puns, asides, unlikely stories and engaging people, inspiring feats and thought-provoking dilemmas. Leaving you armed and ready to decide what you think about one of the decade's big ideas: big data.

From predictive policing to self-surveillance to private security, the potential uses to of big data in crime control pose serious legal and ethical challenges relating to privacy, discrimination, and the presumption of innocence. The book is about the impacts of the use of big data analytics on social and crime control and on fundamental liberties. Drawing on research from Europe and the US, this book identifies the various ways in which law and ethics intersect with the application of big data in social and crime control, considers potential challenges to human rights and democracy and recommends regulatory solutions and best practice. This book focuses on changes in knowledge production and the manifold sites of contemporary surveillance, ranging from self-surveillance to corporate and state surveillance. It tackles the implications of big data and predictive algorithmic analytics for social justice, social equality, and social power: concepts at the very core of crime and social control. This book will be of interest to scholars and students of criminology, sociology, politics and socio-legal studies.

Data Science gets thrown around in the press like it'smagic. Major retailers are predicting everything from when theircustomers are pregnant to when they want a new pair of ChuckTaylors. It's a brave new world where seemingly meaningless datacan be transformed into valuable insight to drive smart businessdecisions. But how does one exactly do data science? Do you have to hireone of these priests of the dark arts, the "data scientist," toextract this gold from your data? Nope. Data science is little more than using straight-forward steps toprocess raw data into actionable insight. And in DataSmart, author and data scientist John Foreman will show you howthat's done within the familiar environment of aspreadsheet. Why a spreadsheet? It's comfortable! You get to look at the dataevery step of the way, building confidence as you learn the tricksof the trade. Plus, spreadsheets are a vendor-neutral place tolearn data science without the hype. But don't let the Excel sheets fool you. This is a book forthose serious about learning the analytic techniques, the math andthe magic, behind big data. Each chapter will cover a different technique in aspreadsheet so you can follow along: Mathematical optimization, including non-linear programming andgenetic algorithms Clustering via k-means, spherical k-means, and graphmodularity Data mining in graphs, such as outlier detection Supervised AI through logistic regression, ensemble models, andbag-of-words models Forecasting, seasonal adjustments, and prediction intervalsthrough monte carlo simulation Moving from spreadsheets into the R programming language You get your hands dirty as you work alongside John through eachtechnique. But never fear, the topics are readily applicable andthe author laces humor throughout. You'll even learnwhat a dead squirrel has to do with optimization modeling, whichyou no doubt are dying to know.

Get the expert perspective and practical advice on big data The Big Data-Driven Business: How to Use Big Data to Win Customers, Beat Competitors, and Boost Profits makes the case that big data is for real, and more than just big hype. The book uses real-life examples—from Nate Silver to Copernicus, and Apple to Blackberry—to demonstrate how the winners of the future will use big data to seek the truth. Written by a marketing journalist and the CEO of a multi-million-dollar B2B marketing platform that reaches more than 90% of the U.S. business population, this book is a comprehensive and accessible guide on how to win customers, beat competitors, and boost the bottom line with big data. The marketplace has entered an era where the customer holds all the cards. With unprecedented choice in both the consumer world and the B2B world, it's imperative that businesses gain a greater understanding of their customers and prospects. Big data is the key to this insight, because it provides a comprehensive view of a company's customers—who they are, and who they may be tomorrow. The Big Data-Driven Business is a complete guide to the future of business as seen through the lens of big data, with expert advice on real-world applications. Learn what big data is, and how it will transform the enterprise Explore why major corporations are betting their companies on marketing technology Read case studies of big data winners and losers Discover how to change privacy and security, and remodel marketing Better information allows for better decisions, better targeting, and better reach. Big data has become an indispensable tool for the most effective marketers in the business, and it's becoming less of a competitive advantage and more like an industry standard. Remaining relevant as the marketplace evolves requires a full understanding and application of big data, and The Big Data-Driven Business provides the practical guidance businesses need.

Written by a leading expert in the field, this account focuses on the convergence of two major trends in information management—big data and information governance—by taking a strategic approach oriented around business cases and industry imperatives. With the advent of new technologies, enterprises are expanding and handling very large volumes of data; this book, nontechnical in nature and geared toward business audiences, encourages the practice of establishing appropriate governance over big data initiatives and addresses how to manage and govern big data, highlighting the relevant processes, procedures, and policies. It teaches readers to understand how big data fits within an overall information governance program; quantify the business value of big data; apply information governance concepts such as stewardship, metadata, and organization structures to

big data; appreciate the wide-ranging business benefits for various industries and job functions; sell the value of big data governance to businesses; and establish step-by-step processes to implement big data governance.

Due to the scale and complexity of data sets currently being collected in areas such as health, transportation, environmental science, engineering, information technology, business and finance, modern quantitative analysts are seeking improved and appropriate computational and statistical methods to explore, model and draw inferences from big data. This book aims to introduce suitable approaches for such endeavours, providing applications and case studies for the purpose of demonstration. Computational and Statistical Methods for Analysing Big Data with Applications starts with an overview of the era of big data. It then goes onto explain the computational and statistical methods which have been commonly applied in the big data revolution. For each of these methods, an example is provided as a guide to its application. Five case studies are presented next, focusing on computer vision with massive training data, spatial data analysis, advanced experimental design methods for big data, big data in clinical medicine, and analysing data collected from mobile devices, respectively. The book concludes with some final thoughts and suggested areas for future research in big data. Advanced computational and statistical methodologies for analysing big data are developed Experimental design methodologies are described and implemented to make the analysis of big data more computationally tractable Case studies are discussed to demonstrate the implementation of the developed methods Five high-impact areas of application are studied: computer vision, geosciences, commerce, healthcare and transportation Computing code/programs are provided where appropriate

Data Science and Big Data Analytics is about harnessing the power of data for new insights. The book covers the breadth of activities and methods and tools that Data Scientists use. The content focuses on concepts, principles and practical applications that are applicable to any industry and technology environment, and the learning is supported and explained with examples that you can replicate using open-source software. This book will help you: Become a contributor on a data science team Deploy a structured lifecycle approach to data analytics problems Apply appropriate analytic techniques and tools to analyzing big data Learn how to tell a compelling story with data to drive business action Prepare for EMC Proven Professional Data Science Certification Corresponding data sets are available from the book's page at Wiley which you can find on the Wiley site by searching for the ISBN 9781118876138. Get started discovering, analyzing, visualizing, and presenting data in a meaningful way today!

Did you know that your business already has the world's greatest information-tracking team working tirelessly for you 24/7 to gather all the info you could possibly need to find your next customers? Between brand tracking, CRM programs, and online behavior tracking, as well as the always-dependable trade shows and satisfaction studies, mounds of marketing metrics are being generated for you across various touchpoints and channels. The numbers available to you are mind-blowing--but the amount itself can be mind-numbing. Where can one begin to filter through it all to find what is most beneficial for their company?Locked in the vast quantity of information are accurate, data-driven answers to every marketing question--and analytic dashboards are the key to finding it all. In It's Not the Size of the Data--It's How You Use It, marketing expert Koen Pauwels introduces readers to these transformative web-based tools that gather, synthesize, and visually display essential data in real time, directly connecting marketing with performance. He then supplies a simple yet rigorous methodology that explains step by step how to:• Gain crucial IT support • Build a rock-solid database • Select key leading performance indicators • Design the optimal dashboard layout • Use marketing analytics to improve decisions and reap rewardsThere is simply too much customer-produced information out there today for marketing teams to go with gut decisions or the same old standbys. Dashboard analytics will bring scientific precision and insight to the marketing efforts of any size organization, in any industry, and turn this eye-popping data into a specific plan of attack.

Exploit the power and potential of Big Data to revolutionize business outcomes Big Data Revolution is a guide to improving performance, making better decisions, and transforming business through the effective use of Big Data. In this collaborative work by an IBM Vice President of Big Data Products and an Oxford Research Fellow, this book presents inside stories that demonstrate the power and potential of Big Data within the business realm. Readers are guided through tried-and-true methodologies for getting more out of data, and using it to the utmost advantage. This book describes the major trends emerging in the field, the pitfalls and triumphs being experienced, and the many considerations surrounding Big Data, all while guiding readers toward better decision making from the perspective of a data scientist. Companies are generating data faster than ever before, and managing that data has become a major challenge. With the right strategy, Big Data can be a powerful tool for creating effective business solutions – but deep understanding is key when applying it to individual business needs. Big Data Revolution provides the insight executives need to incorporate Big Data into a better business strategy, improving outcomes with innovation and efficient use of technology. Examine the major emerging patterns in Big Data Consider the debate surrounding the ethical use of data Recognize patterns and improve personal and organizational performance Make more informed decisions with quantifiable results In an information society, it is becoming increasingly important to make sense of data in an economically viable way. It can drive new revenue streams and give companies a competitive advantage, providing a way forward for businesses navigating an increasingly complex marketplace. Big Data Revolution provides expert insight on the tool that can revolutionize industries.

The data lake is a daring new approach for harnessing the power of big data technology and providing convenient self-service capabilities. But is it right for your company? This book is based on discussions with practitioners and executives from more than a hundred organizations, ranging from data-driven companies such as Google, LinkedIn, and Facebook, to governments and traditional corporate enterprises. You'll learn what a data lake is, why enterprises need one, and how to build one successfully with the best practices in this book. Alex Gorelik, CTO and founder of Waterline Data, explains why old systems and processes can no longer support data needs in the enterprise. Then, in a collection of essays about data lake implementation, you'll examine data lake initiatives, analytic projects, experiences, and best practices from data experts working in various industries. Get a succinct introduction to data warehousing, big data, and

data science Learn various paths enterprises take to build a data lake Explore how to build a self-service model and best practices for providing analysts access to the data Use different methods for architecting your data lake Discover ways to implement a data lake from experts in different industries

Practical data design tips from a data visualization expert ofthe modern age Data doesn?t decrease; it is ever-increasing and can beoverwhelming to organize in a way that makes sense to its intendedaudience. Wouldn?t it be wonderful if we could actually visualizedata in such a way that we could maximize its potential and tell astory in a clear, concise manner? Thanks to the creative genius ofNathan Yau, we can. With this full-color book, data visualizationguru and author Nathan Yau uses step-by-step tutorials to show youhow to visualize and tell stories with data. He explains how togather, parse, and format data and then design high qualitygraphics that help you explore and present patterns, outliers, andrelationships. Presents a unique approach to visualizing and telling storieswith data, from a data visualization expert and the creator offlowingdata.com, Nathan Yau Offers step-by-step tutorials and practical design tips forcreating statistical graphics, geographical maps, and informationdesign to find meaning in the numbers Details tools that can be used to visualize data-nativegraphics for the Web, such as ActionScript, Flash libraries, PHP,and JavaScript and tools to design graphics for print, such as Rand Illustrator Contains numerous examples and descriptions of patterns andoutliers and explains how to show them Visualize This demonstrates how to explain data visually so thatyou can present your information in a way that is easy tounderstand and appealing.

How a new understanding of warfare can help the military fight today’s conflicts more effectively The way wars are fought has changed starkly over the past sixty years. International military campaigns used to play out between armies at central fronts. Today's conflicts find major powers facing rebel insurgencies deploying elusive methods, from improvised explosives to terrorist attacks. Presenting a transformative understanding of these contemporary confrontations, Small Wars, Big Data shows that a revolution in the study of conflict yields new insights into terrorism, civil wars, and foreign interventions. Modern warfare is not about struggles over territory but over people; civilians—and the information they might provide—can turn the tide at critical junctures. Drawing lessons from conflicts in locations around the world, Small Wars, Big Data provides groundbreaking perspectives for how small wars can be better strategized and favorably won.

Every company has growing pains. When you are the owner or senior management of a business, just running your business can keep you from focusing on its growth potential and taking the steps to get there. In Size Does Matter: Grow Your Business, Thinking Big, Rony Pawar takes you on a step-by-step journey to set your business up for success and growth. Rony consistently reminds you that no business can stand still. To achieve growth, your business needs a strategic plan, one that focuses on continuous improvement. Right from the first chapter, Rony will start you on your journey of creating a strategic plan and putting your business on the path to success! Rony harnesses his knowledge and experience gained from working with large and small businesses around the globe, to provide you an action plan that starts with your leadership and addresses several key areas, including management, building revenue, the customer experience, and innovation. Throughout the book, Rony takes these concepts and provides real-world examples to help you see the potential of your business to reach the next level. It is a guide for your journey in the world of business growth. Whether you are just starting out in your business or have been at it for years, Rony's energizing and personal approach will help you see your business in a whole new light and get you motivated to grow!

Practical Data Science for Information Professionals provides an accessible introduction to a potentially complex field, providing readers with an overview of data science and a framework for its application. It provides detailed examples and analysis on real data sets to explore the basics of the subject in three principle areas: clustering and social network analysis; predictions and forecasts; and text analysis and mining. As well as highlighting a wealth of user-friendly data science tools, the book also includes some example code in two of the most popular programming languages (R and Python) to demonstrate the ease with which the information professional can move beyond the graphical user interface and achieve significant analysis with just a few lines of code. After reading, readers will understand: · the growing importance of data science · the role of the information professional in data science · some of the most important tools and methods that information professionals can use. Bringing together the growing importance of data science and the increasing role of information professionals in the management and use of data, Practical Data Science for Information Professionals will provide a practical introduction to the topic specifically designed for the information community. It will appeal to librarians and information professionals all around the world, from large academic libraries to small research libraries. By focusing on the application of open source software, it aims to reduce barriers for readers to use the lessons learned within.

Tips, techniques, and trends on how to use dashboard technology tooptimize business performance Business performance management is a hot new management disciplinethat delivers tremendous value when supported by informationtechnology. Through case studies and industry research, this bookshows how leading companies are using performance dashboards toexecute strategy, optimize business processes, and improveperformance. Wayne W. Eckerson (Hingham, MA) is the Director of Research for TheData Warehousing Institute (TDWI), the leading association ofbusiness intelligence and data warehousing professionals worldwidethat provide high-quality, in-depth education, training, andresearch. He is a columnist for SearchCIO.com, DM Review,Application Development Trends, the Business Intelligence Journal,and TDWI Case Studies & Solution.

If you want to outsmart a crook, learn his tricks—Darrell Huff explains exactly how in the classic How to Lie with Statistics. From distorted graphs and biased samples to misleading averages, there are countless statistical dodges that lend cover to anyone with an ax to grind or a product to sell. With abundant examples and illustrations, Darrell Huff’s lively and engaging primer clarifies the basic principles of statistics and explains how they’re used to present information in honest and not-so-honest ways. Now even more indispensable in our data-driven world than it was when first published, How to Lie with Statistics is the book that generations of readers have relied on to keep from being fooled.

This book presents current progress on challenges related to Big Data management by focusing on the particular challenges associated with context-aware data-intensive applications and services. The book is a state-of-the-art reference discussing progress made, as well as prompting future directions on the theories, practices, standards and strategies that are related to the emerging computational technologies and their association with supporting the Internet of Things advanced functioning for organizational settings including both business and e-science. Apart from inter-operable and inter-cooperative aspects, the book deals with a notable opportunity namely, the current trend in which a collectively shared and generated content is emerged from

Internet end-users. Specifically, the book presents advances on managing and exploiting the vast size of data generated from within the smart environment (i.e. smart cities) towards an integrated, collective intelligence approach. The book also presents methods and practices to improve large storage infrastructures in response to increasing demands of the data intensive applications. The book contains 19 self-contained chapters that were very carefully selected based on peer review by at least two expert and independent reviewers and is organized into the three sections reflecting the general themes of interest to the IoT and Big Data communities: Section I: Foundations and Principles Section II: Advanced Models and Architectures Section III: Advanced Applications and Future Trends The book is intended for researchers interested in joining interdisciplinary and transdisciplinary works in the areas of Smart Environments, Internet of Things and various computational technologies for the purpose of an integrated collective computational intelligence approach into the Big Data era.

From the gas-guzzling, eight-passenger SUVs, to sprawling suburban mansions, to Super Big Gulps, to human anatomy - have you wondered if bigger is better? Why are some people obsessed with size? You will find the answer in Does Size Matter?, from the popular F.Y.I. book series, along with answers to more than 150 other questions that stretch the bounds of curiosity. Does Size Matter? is a smart and authoritative book of useful information on myriad topics, ranging from body science to weird science, the animal kingdom to earth and space, love and lust to origins and traditions, and people to sports. Our writers demystify urban legends and inform on everything from the strange to the sublime.

Big data raises more questions than it answers, particularly for those organizations struggling to deal with what has become an overwhelming deluge of data. It can offer marketers more than simple tactical predictive analytics, but organizations need a bigger picture, one that generates some real insight into human behaviour, to drive consumer strategy rather than just better targeting techniques. Humanizing Big Data guides marketing managers, brand managers, strategists and senior executives on how to use big data strategically to redefine customer relationships for better customer engagement and an improved bottom line. Humanizing Big Data provides a detailed understanding of the way to approach and think about the challenges and opportunities of big data, enabling any brand to realize the value of their current and future data assets. First it explores the 'nuts and bolts' of data analytics and the way in which the current big data agenda is in danger of losing credibility by paying insufficient attention to what are often fundamental tenets in any form of analysis. Next it sets out a manifesto for a smart data approach, drawing on an intelligent and big picture view of data analytics that addresses the strategic business challenges that businesses face. Finally it explores the way in which datafication is changing the nature of the relationship between brands and consumers and why this calls for new forms of analytics to support rapidly emerging new business models. After reading this book, any brand should be in a position to make a step change in the value they derive from their data assets.

Big Data is the biggest game-changing opportunity for marketing and sales since the Internet went mainstream almost 20 years ago. The data big bang has unleashed torrents of terabytes about everything from customer behaviors to weather patterns to demographic consumer shifts in emerging markets. This collection of articles, videos, interviews, and slideshares highlights the most important lessons for companies looking to turn data into above-market growth: Using analytics to identify valuable business opportunities from the data to drive decisions and improve marketing return on investment (MROI) Turning those insights into well-designed products and offers that delight customers Delivering those products and offers effectively to the marketplace.The goldmine of data represents a pivot-point moment for marketing and sales leaders. Companies that inject big data and analytics into their operations show productivity rates and profitability that are 5 percent to 6 percent higher than those of their peers. That's an advantage no company can afford to ignore.

Big DataDoes Size Matter?Bloomsbury Publishing