

Data Scientist The Definitive Guide To Becoming A Data Scientist English Edition

Develop the skills you need to run machine learning workloads in Azure and pass the DP-100 exam with ease
Key Features>Create end-to-end machine learning training pipelines, with or without codeTrack experiment progress using the cloud-based MLflow-compatible process of Azure ML servicesOperationalize your machine learning models by creating batch and real-time endpointsBook Description The Azure Data Scientist Associate Certification Guide helps you acquire practical knowledge for machine learning experimentation on Azure. It covers everything you need to pass the DP-100 exam and become a certified Azure Data Scientist Associate. Starting with an introduction to data science, you'll learn the terminology that will be used throughout the book and then move on to the Azure Machine Learning (Azure ML) workspace. You'll discover the studio interface and manage various components, such as data stores and compute clusters. Next, the book focuses on no-code and low-code experimentation, and shows you how to use the Automated ML wizard to create and deploy your datasets. You'll learn how to run end-to-end machine learning designs using the designer provided in Azure ML Studio. You'll then explore the Azure ML Software Development Kit (SDK) for Python and advance to creating experiment and publishing models using code. The book also guides you in optimizing your model's hyperparameters using Hyperdrive before demonstrating how to use responsible AI tools to interpret and debug your models. Once you have a trained model, you'll learn to operationalize it for batch or real-time inferences and monitor it in production. By the end of this Azure certification study guide, you'll have gained the knowledge and the practical skills required to pass the DP-100 exam. What you will learn>Create a working environment for data science workloads on AzureRun data experiments using Azure Machine Learning servicesCreate training and inference pipelines using the designer or codeDiscover the best model for your dataset using Automated MLUse hyperparameter tuning to optimize trained modelsDeploy, use, and monitor models in productionInterpret the predictions of a trained modelWho this book is for This book is for developers who want to infuse their applications with AI capabilities and data scientists looking to scale their machine learning experiments in the Azure cloud. Basic knowledge of Python is needed to follow the code samples used in the book. Some experience in training machine learning models in Python using common frameworks like scikit-learn will help you understand the content more easily. Scientific progress depends on good research, and good research needs good statistics. But statistical analysis is tricky to get right, even for the best and brightest of us. You'd be surprised how many scientists are doing it wrong. Statistics Done Wrong is a pithy, essential guide to statistical blunders in modern science that will show you how to keep your research blunder-free. You'll examine embarrassing errors and omissions in recent research, learn about the misconceptions and scientific politics that allow these mistakes to happen, and begin your quest to reform the way you and your peers do statistics. You'll find advice on: -Asking the right question, designing the right experiment, choosing the right statistical analysis, and sticking to the plan -How to think about statistics, insignificance, confidence intervals, and regression -Choosing the right sample size and avoiding false positives -Reporting your analysis and publishing your data and source code -Procedures to follow, precautions to take, and analytical software that can help Scientists: Read this concise, powerful guide to help you produce statistically sound research. Statisticians: Give this book to everyone you know. The first step toward statistics done right is Statistics Done Wrong.

The Ultimate Guide to Data Analytics, Data Mining, Data Warehousing, Data Analytics, Modelling, Visualization, Regression Analysis
A Definitive Guide to Data Mining, Data Warehousing, Data Analytics, Modelling, Visualization, Regression Analysis
Azure Data Scientist Associate Certification Guide
Data ScientistThe Definitive Guide to Becoming a Data ScientistTechnics Publications
Learn how to build a data science technology stack and perform good data science with repeatable methods. You will learn how to turn data lakes into business assets. The data science technology stack demonstrated in Practical Data Science is built from components in general use in the industry. Data scientist Andreas Vermeulen demonstrates in detail how to build and provision a technology stack to yield repeatable results. He shows you how to apply practical methods to extract actionable business knowledge from data lakes consisting of data from a polyglot of data types and dimensions. What You'll Learn Become fluent in the essential concepts and terminology of data science and data engineering Build and use a technology stack that meets industry criteria Master the methods for retrieving actionable business knowledge Coordinate the handling of polyglot data types in a data lake for repeatable results Who This Book Is For Data scientists and data engineers who are required to convert data from a data lake into actionable knowledge for their business, and students who aspire to be data scientists and data engineers
R for Political Data Science: A Practical Guide is a handbook for political scientists new to R who want to learn the most useful and common ways to interpret and analyze political data. It was written by political scientists, thinking about the many real-world problems faced in their work. The book has 16 chapters and is organized in three sections. The first, on the use of R, is for those users who are learning R or are migrating from another software. The second section, on econometric models, covers OLS, binary and survival models, panel data, and causal inference. The third section is a data science toolbox of some the most useful tools in the discipline: data imputation, fuzzy merge of large datasets, web mining, quantitative text analysis, network analysis, mapping, spatial cluster analysis, and principal component analysis. Key features: Each chapter has the most up-to-date and simple option available for each task, assuming minimal prerequisites and no previous experience in R Makes extensive use of the Tidyverse, the group of packages that has revolutionized the use of R Provides a step-by-step guide that you can replicate using your own data Includes exercises in every chapter for course use or self-study Focuses on practical-based approaches to statistical inference rather than mathematical formulae Supplemented by an R package, including all data As the title suggests, this book is highly applied in nature, and is designed as a toolbox for the reader. It can be used in methods and data science courses, at both the undergraduate and graduate levels. It will be equally useful for a university student pursuing a PhD, political consultants, or a public official, all of whom need to transform their datasets into substantive and easily interpretable conclusions.

Are algorithms friend or foe? The human mind is evolutionarily designed to take shortcuts in order to survive. We jump to conclusions because our brains want to keep us safe. A majority of our biases work in our favor, such as when we feel a car speeding in our direction is dangerous and we instantly move, or when we decide not take a bite of food that appears to have gone bad. However, inherent bias negatively affects work environments and the decision-making surrounding our communities. While the creation of algorithms and machine learning attempts to eliminate bias, they are, after all, created by human beings, and thus are susceptible to what we call algorithmic bias. In Understand, Manage, and Prevent Algorithmic Bias, author Tobias Baer helps you understand where algorithmic bias comes from, how to manage it as a business user or regulator, and how data science can prevent bias from entering statistical algorithms. Baer expertly addresses some of the 100+ varieties of natural bias such as confirmation bias, stability bias, pattern-recognition bias, and many others. Algorithmic bias mirrors—and originates in—these human tendencies. Baer dives into topics as diverse as anomaly detection, hybrid model structures, and self-improving machine learning. While most writings on algorithmic bias focus on the dangers, the core of this positive, fun book points toward a path where bias is kept at bay and even eliminated. You'll come away with managerial techniques to develop unbiased algorithms, the ability to detect bias more quickly, and knowledge to create unbiased data. Understand, Manage, and Prevent Algorithmic Bias is an innovative, timely, and important book that belongs on your shelf. Whether you are a seasoned business executive, a data scientist, or simply an enthusiast, now is a crucial time to be educated about the impact of algorithmic bias on society and take an active role in fighting bias. What You'll Learn Study the many sources of algorithmic bias, including cognitive biases in the real world, biased data, and statistical artifact Understand the risks of algorithmic biases, how to detect them, and managerial techniques to prevent or manage them Appreciate how machine learning both introduces new sources of algorithmic bias and can be a part of a solutionBe familiar with specific statistical techniques a data scientist can use to detect and overcome algorithmic bias Who This Book Is For Business executives of companies using algorithms in daily operations; data scientists (from students to seasoned practitioners) developing algorithms; compliance officials concerned about algorithmic bias; politicians, journalists, and philosophers thinking about algorithmic bias in terms of its impact on society and possible regulatory responses; and consumers concerned about how they might be affected by algorithmic bias

Foundations of Data Science

A Guide to Data Science for Fraud Detection

Data Science From Scratch

Data Science

The Ultimate Guide to Data Analytics, Data Mining, Data Warehousing, Data Visualization, Regression Analysis, Database Querying, Big Data for Business and Machine Learning for Beginners

The Decision Maker's Handbook to Data Science

Import, Tidy, Transform, Visualize, and Model Data

We're living in a digital world. Most of our global economy is digital and the sheer volume of data is stupendous. It's 2020 and we're living in the future. Data Scientist is one of the hottest job on the market right now. Demand for data science is huge and will only grow, and it seems like it will grow much faster than the actual number of data scientists. So if you want to make a career change and become a data scientist, now is the time. This book will guide you through the process. From my experience of working with multiple companies as a project manager, a data science consultant or a CTO, I was able to see the process of hiring data scientists and building data science teams. I know what's important to land your first job as a data scientist, what skills you should acquire, what you should show during a job interview. 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 into 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.

Data science comprises one of the most promising and in-demand career paths for skilled professionals. Today, successful data professionals understand that they must advance past the traditional skills of analyzing large amounts of data, data mining, and programming skills. In order to uncover useful intelligence for their organizations, data scientists must master the full spectrum of the data science life cycle and possess a level of flexibility and understanding to maximize returns at each phase of the process. There are many resources and courses online that you can use to learn more about data science, but with so much information available, it can become overwhelming. One of the best ways to learn about data science is to understand different machine learning concepts, statistics, and artificial intelligence to help you design models to perform an analysis. This book has all the information you need to learn what data science is, and what the prerequisites are to become a data scientist. In this book, you will learn what data science is about. Discover the difference between data science and business intelligence. Explore the tools required for data science. Find out the technical and non-technical skills every data scientist must have. Figure out how to create a visualization of the data set with clear and easy examples. Get advice on developing a Predictive Model Using R. Uncover detailed applications of data science. And much more! Buy it right now!

The only how-to guide offering a unified, systemic approach to acquiring, cleaning, and managing data in R Every experienced practitioner knows that preparing data for modeling is a painstaking, time-consuming process. Adding to the difficulty is that most modelers learn the steps involved in cleaning and managing data piecemeal, often on the fly, or they develop their own ad hoc methods. This book helps simplify their task by providing a unified, systematic approach to acquiring, modeling, manipulating, cleaning, and maintaining data in R. Starting with the very basics, data scientists Samuel E. Buttrety and Lynn R. Whitaker walk readers through the entire process. From what data looks like and what it should look like, they progress through all the steps involved in getting data ready for modeling. They describe best practices for acquiring data from numerous sources; explore key issues in data modeling and calculations; big data; parallel processing; merging, matching, and checking for duplicates; and outlining highly efficient and reliable techniques for documenting data and recordkeeping, including audit trails, getting data back out of R, and more. The only single-source guide to R data and its preparation, it describes best practices for acquiring, manipulating, cleaning, and maintaining data. Begins with the basics and walks readers through all the steps necessary to get data ready for the modeling process Provides expert guidance on how to document the processes described so that they are reproducible Written by seasoned professionals, it provides both introductory and advanced techniques Features case studies with supporting data and R code, hosted on a companion website& A Data Scientist's Guide to Acquiring, Cleaning and Managing Data in R is a valuable working resource/bench manual for practitioners who collect and analyze data, lab scientists and research associates of all levels of experience, and graduate-level data mining students.

Thinking Clearly with Data

End-to-End Data Science with SAS

A Hands-On Programming Guide

How to Become a Data Scientist. Complete Guide to Learning the Data Science Process & what YOU NEED to Know About: Analytics, Mining, Structures, Management, Driven, Privacy

Big Data Processing Made Simple

A Guide for Non-Technical Executives, Managers, and Founders

50 Essential Concepts

Hiring Data Scientists and Machine Learning Engineers is a concise, practical guide to help you hire the right people for your organization. The book will help you navigate the plethora of data science related roles and skills and help you create an effective hiring strategy to suit your organization's needs.

Every major company has or will soon have a data science program. Most fall, expensively, imperiling their executive sponsors. Unfortunately, executives have been misled to focus on the latest buzzwords. Although buzzwords change—big data, data science, machine learning, deep learning, and artificial intelligence—the distraction from fundamentals manifests as a predictable trajectory from exuberant program launch, to stagnation, to awkward decommissioning. After architecting data science programs at over a dozen companies, the author provides a reliable framework for successful data science programs. Surprisingly, software and algorithms are secondary. Rather, the key is understanding how the available data aligns to the problem to be solved. The business executive understands the problem sufficiently to enforce this alignment, while data scientists act on it. But executives tend to underestimate their role and thereby fail to construct the necessary connective tissue with their data scientists. This book provides business executives with a concrete exercise, populating a "Master Table," accessible to nontechnical managers and data scientists, which serves as the connective tissue between them. Rather than teach a diluted version of data science, this book describes how to start projects and how to detect and fix problems—the moments when leadership is critical. Insights are provided through real world examples, including a Playbook featuring common projects. The intended audience is executives (C-suite through VP).

However, ambitious mid-level managers and data scientists will also benefit. A field guide for the unique challenges of data science leadership, filled with transformative insights, personal experiences, and industry examples. In How To Lead in Data Science you will learn: Best practices for leading projects while balancing complex trade-offs Specifying, prioritizing, and planning projects from vague requirements Navigating structural challenges in your organization Working through project failures with positivity and tenacity Growing your team with coaching, mentoring, and advising Crafting technology roadmaps and championing successful projects Driving diversity, inclusion, and belonging within teams Archetype a long-term business strategy and data roadmap as an executive Delivering a data-driven culture and structuring productive data science organizations How to Lead in Data Science is full of techniques for leading data science at every seniority level—from heading up a single project to overseeing a whole company's data strategy. Authors Jake Chong and Yue Cathy Chang share hard-won advice that they've developed building data teams for LinkedIn, Acorns, Yiren Digital, large asset-management firms, Fortune 50 companies, and more. You'll find advice on plotting your long-term career advancement, as well as quick wins you can put into practice right away. Carefully crafted assessments and interview scenarios encourage introspection, reveal personal blind spots, and highlight development areas. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Lead in Data Science and scales. Dr. Elser has formulated a reliable framework for successful data science programs. Surprisingly, software and algorithms are secondary. Rather, the key is understanding how the available data aligns to the problem to be solved. The business executive understands the problem sufficiently to enforce this alignment, while data scientists act on it. But executives tend to underestimate their role and thereby fail to construct the necessary connective tissue with their data scientists. This book provides business executives with a concrete exercise, populating a "Master Table," accessible to nontechnical managers and data scientists, which serves as the connective tissue between them. Rather than teach a diluted version of data science, this book describes how to start projects and how to detect and fix problems—the moments when leadership is critical. Insights are provided through real world examples, including a Playbook featuring common projects. The intended audience is executives (C-suite through VP). It's filled with best practices for linking up and producing exceptional results, even when beginning with vague requirements or unclear expectations. You'll find a clearly presented modern leadership framework based on current case studies, with insights reaching all the way to Aristotle and Confucius. As you read, you'll build practical skills to grow and improve your team, your company's data culture, and yourself. What's inside How to coach and mentor team members Navigate an organization's structural challenges Secure commitments from other teams and partners Stay current with the technology landscape Advance your career About the reader For data science practitioners at all levels. About the author Dr. Jake Chong and Yue Cathy Chang build, lead, and grow high-performing data teams across industries in public and private companies, such as Acorns, LinkedIn, large asset-management firms, and Fortune 50 companies. Table of Contents 1 What makes a successful data scientist? PART 1 THE TECH LEAD: CULTIVATING LEADERSHIP 2 Capabilities for leading projects 3 Virtues for leading projects PART 2 THE MANAGER: NURTURING A TEAM 4 Capabilities for leading people 5 Virtues for leading people PART 3 THE DIRECTOR: GOVERNING A FUNCTION 6 Capabilities for leading a function 7 Virtues for leading a function PART 4 THE EXECUTIVE: INSPIRING AN INDUSTRY 8 Capabilities for leading a company 9 Virtues for leading a company PART 5 THE LOOP AND THE FUTURE 10 Landscape, organization, opportunity, and practice 11 Leading in data science and a future outlook

An engaging introduction to data science that emphasizes critical thinking over statistical techniques An introduction to data science or statistics shouldn't involve proving complex theorems or memorizing obscure terms and formulas, but that is exactly what most introductory quantitative textbooks emphasize. In contrast, Thinking Clearly with Data focuses, first and foremost, on critical thinking and conceptual understanding in order to teach students how to be better consumers and analysts of the kinds of quantitative information that they encounter in the world and, if so, whether or not it's useful. Instead of the usual how-to make the most informative comparison or answering question, what questions to ask others who are making arguments, what statistics are particularly informative, misleading, how quantitative evidence should and shouldn't be used, and how to make better decisions by using moral values as well as data. Filled with real-world examples, the book shows how its thinking tools apply to problems in a wide variety of subjects, including elections, civil conflict, crime, terrorism, financial crises, health care, sports, music, and space travel. Above all else, Thinking Clearly with Data demonstrates why, despite the many benefits of our data-driven age, data can never be a substitute for thinking. An ideal textbook for introductory quantitative methods courses in data science, statistics, political science, economics, psychology, sociology, public policy, and other fields Introduces the basic toolkit of data analysis—including sampling, hypothesis testing, Bayesian inference, regression, experiments, instrumental variables, differences in sports, and regression discontinuity Uses real-world examples and data from a wide variety of subjects Includes practice questions and data exercises

R for Data Science

A Guide for the Data Scientist

Python for Data Analysis & Analytics

Learn Data Science from Scratch

How to Talk to Data Scientists

Statistics Done Wrong

Data Scientist

55% OFF for Bookstores! NOW at \$ 34.95 Instead of \$ 54.19 Are You Looking For The Best Beginners Guide To Discovering Data Analysis And Analytics With Python?Do You Want To Enter The World Of Data Science And How To Leverage Python For It?Do You Want To Get A Thorough Introduction To Machine Learning? If yes, then this Guide is for you! This is the Top Guide to learning Data Analysis & Analytics. Talking about the IT world, there are many options when you have to choose language programming to learn and then to use for developing your career, especially if you want to become a Data Scientist. This Handbook will not only give you reasons on why you need to learn data science, but it will also tell you why learning data science with Python training is the better option. In this book you will: Have a Clear and Exhaustive Explanation About Data Analysis and Why It Is So Important Today in the Business World, organizations of all sizes rely on the insights they extract from the data they have to measure progress, make informed decisions, plan for the future, and so on. Data scientists are the people who process and organize the data with scientific methods, algorithms, and other techniques. Understand Why Python is Preferred to Use For Data Analysis Over Other Tools and the reasons why all the benefits of using Python made it the best tool to learn data science. Learn How to Carry Out Work More and More Complex and Difficult to be updated on new themes and trends in the sector and carry out small independent jobs to finance your projects ...& Lot More! Your Customers will never stop to use this book. Are you completely new to programming and want to learn how to code, but don't know where to begin? Are you looking to upgrade your data wrangling skills to future-proof your career and break into Data Science and Analytics? Python is one of the most valuable and interesting languages for data science. Therefore, the popularity of Python is growing day by day, especially in the world of data analysis or data sciences. This Definitive Guide will combine Data Analysis and Python to help your customer build amazing products and help businesses Buy it NOW and let your customers get addicted to this amazing book!

Learn how to use R to turn raw data into insight, knowledge, and understanding. 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, R for Data Science is designed to get you doing data science as quickly as possible. Authors Hadley Wickham and Garrett Grolemund guide you through the steps of importing, wrangling, exploring, and modeling your data and communicating the results. You'll get a complete, big-picture understanding of the data science cycle, along with basic tools you need to manage the details. Each section of the book is paired with exercises to help you practice what you've learned along the way. You'll learn how to: Wrangle—transform your datasets into a form convenient for analysis Program—learn powerful R tools for solving data problems with greater clarity and ease Explain—examine your data, generate hypotheses, and quickly test them Model—provide a low-dimensional summary that captures true "signals" in your dataset Communicate—learn R Markdown for integrating prose, code, and results

55% OFF for Bookstores! NOW at \$ 31.97 instead of \$ 41.97! Do you know that according to Harvard Business Review the Data Scientist is the sexiest job of the 21st century? And for a reason! Your Customers Will Love This Amazing Step-By-Step Guide! If "sexy" means having rare qualities that are much in demand, data scientists are already there. They are expensive to hire and, given the very competitive market for their services, difficult to retain. There simply aren't a lot of people with their combination of scientific background and computational and analytical skills. This guidebook will start off by taking a look at what Data Science is all about. There are a lot of companies throughout many industries that are already working with Data Science because they see the enormous value that they can get. We will take a look at some of the steps that can help with the Data Science lifecycle and how companies are able to benefit from implementing it for themselves. If you have spent any time in business recently, then it is likely that you are familiar with the concept of data-driven solution. Many companies are jumping on board and learning more about how to implement Data Science in their business. And with the help of this guidebook, you will be able to do some of the same as well. When you are ready to jump on board and learn more about Data Science and how you and your business can benefit from this process, make sure to check out this guidebook to get started. In "Learn Data Science from Scratch" you will discover: What Data Science is, and What are its Concepts & Terms, so that you can have full knowledge on the subject, and you can be sure to easily understand and learn every argument and skill related to this topic How to Master Python in just a few weeks through practical examples and applications and simple explanations Practical Codes and Exercises to Use Python, that you can easily use to practice your skills and that you can test and tweak to your heart's content What are the Best, Essential Libraries for Data Science in Python, so you can have access to the best and most-useful functions without the need of writing specific codes from scratch The Topnotch Data Mining Techniques in Data Science, to Always get the most out of your data gathering process, thanks to a complete chapter that will show you what are the best ways to maximize the value of data investments The Data Science Pipeline This book is essential for anyone who wants to study Data Science and learn how the world is moving to an open-source platform. Whether you are a software engineer or a project manager, jump to the next level by developing a data-driven approach and learning how to define a data-specific vision of your business! Order Tour Copy NOW and Let Your Customers Get Addicted to This Amazing Book!

Summary You are going to need more than technical knowledge to succeed as a data scientist. Build a Career in Data Science teaches you what school leaves out, from how to land your first job to the lifecycle of a data science project, and even how to become a manager. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology What are the keys to a data scientist's long-term success? Blending your technical know-how with the right "soft skills" turns out to be a central ingredient of a rewarding career. About the book Build a Career in Data Science is your guide to landing your first data science job and developing into a valued senior employee. By following clear and simple instructions, you'll learn to craft an amazing resume and ace your interviews. In this demanding, rapidly changing field, it can be challenging to keep projects on track, adapt to company needs, and manage tricky stakeholders. You'll love the insights on how to handle expectations, deal with failures, and plan your career path in the stories from seasoned data scientists included in the book. What's inside Creating a portfolio of data science projects Assessing and negotiating an offer Leaving gracefully and moving up the ladder Interviews with professional data scientists About the reader For readers who want to begin or advance a data science career. About the author Emily Robinson is a data scientist at Warby Parker. Jacqueline Nolis is a data science consultant and mentor. Table of Contents: PART 1 - GETTING STARTED WITH DATA SCIENCE 1. What is data science? 2. Data science companies 3. Getting the skills 4. Building a portfolio PART 2 - FINDING YOUR DATA SCIENCE JOB 5. The project: Identifying the right job for you 6. The application: R ésumé e s and cover letters 7. The interview: What to expect and how to handle it 8. The offer: Readers who want to accept PART 3 - SETTLING INTO DATA SCIENCE 9. The first months on the job 10. Making an effective analysis 11. Deploying a model into production 12. Working with stakeholders PART 4 - GROWING IN YOUR DATA SCIENCE ROLE 13. When your data science project fails 14. Joining the data science community 15. Leaving your job gracefully 16. Moving up the ladder

A Guide to Quantitative Reasoning and Analysis

Predictive Modelling Concepts Explained in Simple Terms for Beginners

R for Political Data Science

Introduction to Machine Learning with Python

A Practical Guide

The Art of Data Science

Becoming a Data Scientist

Data science is expanding across industries at a rapid pace, and the companies first to adopt best practices will gain a significant advantage. To reap the benefits, decision makers need to have a confident understanding of data science and its application in their organization. It is easy for novices to the subject to feel paralyzed by intimidating buzzwords, but what many don't realize is that data science is in fact quite multidisciplinary—useful in the hands of business analysts, communications strategists, designers, and more. With the second edition of The Decision Maker's Handbook to Data Science, you will learn how to think like a veteran data scientist and approach solutions to business problems in an entirely new way. Author Strylanc Kampakis provides you with the expertise and tools required to develop a solid data strategy that is continuously effective. Ethics and legal issues surrounding data collection and algorithmic bias are some common pitfalls that Kampakis helps you avoid, while guiding you on the path to build a thriving data science culture at your organization. This updated and revised second edition, includes plenty of case studies, tools for project assessment, and expanded content for hiring and managing data scientists Data science is a language that everyone at a modern company should understand across departments. Friction in communication arises most often when management does not connect with what a data scientist is doing or how impactful data collection and storage can be for their organization. The Decision Maker's Handbook to Data Science bridges this gap and readies you for both the present and future of this engaging, comprehensive guide. What You Will Learn Understand how data science can be used within your business. Recognize the differences between AI, machine learning, and statistics. Become skilled at thinking like a data scientist, without being one. Discover how to hire and manage data scientists. Comprehend how to build the right environment in order to make your organization data-driven. Who This Book Is For Starting founders, product managers, higher level managers, and any other non-technical decision makers who are thinking to implement data science in their organization and hire data scientists. A secondary audience includes people looking for a soft introduction into the subject of data science.

This book provides an intuitive and accessible foundation 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 in an accessible, intuitive and accessible manner. The book is suitable for both undergraduate and graduate courses in the design and analysis of algorithms for data.

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 summarize, 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.

Learn what it takes to succeed in the most in-demand tech job Harvard Business Review calls it the sexiest tech job of the 21st century: Data scientists are in demand, and this unique book shows you exactly what employers want and the skill set that separates the quality data scientists from other talented IT professionals. Data science involves extracting, creating, and processing data to turn it into business value. With over 15 years of big data, predictive modeling, and business analytics experience, author Vincent Granville is no stranger to data science. In this one-of-a-kind guide, he provides insight into the essential data science skills, such as statistics and visualization techniques, and covers everything from automated data science risks to common job interview questions, sample resumes, and source code. The applications are endless and varied, automatically detecting spam and phishing, optimizing bid prices in keyword advertising, identifying new molecules to fight cancer, assessing the risk of meteorite impact. Complete with case studies, this book is a must, whether you're looking to become a data scientist or hire one. Explains the four points of data science, the required skills, and how to acquire them, including analytical respces, standard rules, source code, and a dictionary of terms Shows what companies are looking for and how the growing importance of big data has increased the demand for data scientists Features job interview questions, sample resumes, salary surveys, and examples of job ads Case studies explore how data science is used on Wall Street, in bomb detection, for online advertising, and in many other business-critical situations Developing Analytic Talent: Becoming a Data Scientist is essential reading for those aspiring to this hot career choice and for employers seeking the best candidates.

The Ultimate Definitive Guide to Learn Data Science and Coding With Python. Master The Basics of Machine Learning, or Clean Code and Improve Artificial Intelligence

Introduction to Data Visualization and Storytelling

A Definitive Guide on How to Work with Data in Python Programming Language

A Guide to Building the Technology Stack for Turning Data Lakes into Business Assets

Build a Career in Data Science

Spark: The Definitive Guide

The Definitive Guide to Becoming a Data Scientist

"This book describes the process of analyzing data. The authors have extensive experience both managing data analysts and conducting their own data analyses, and this book is a distillation of their experience in a format that is applicable to both practitioners and managers in data science."--Leanpub.com.

Statistical methods are a key part of data science, yet very few data scientists have any formal statistics training. Courses and books on basic statistics rarely cover the topic from a data science perspective. This practical guide explains how to apply various statistical methods to data science, tells you how to avoid their misuse, and gives you advice on what's important and what's not. Many data science resources incorporate statistical methods but lack a deeper statistical programming language, and have some exposure to statistics, this quick reference bridges the gap in an accessible, readable format. With this book, you'll learn: Why exploratory data analysis is a key preliminary step in data science How random sampling can reduce bias and yield a higher quality dataset, even with big data How the principles of experimental design yield definitive answers to questions How to use regression to estimate outcomes and detect anomalies Key classification categories a record belongs to Statistical machine learning methods that "learn" from data Unsupervised learning methods for extracting meaning from unlabeled data

2 comprehensive manuscripts in 1 book Data Science: What the Best Data Scientists Know About Data Analytics, Data Mining, Statistics, Machine Learning, and Big Data - That You Don't Data Science for Business: Predictive Modeling, Data Mining, Data Analytics, Data Warehousing, Data Visualization, Regression Analysis, Database Querying

As our society transforms into a data-driven one, the role of the Data Scientist is becoming more and more important. If you want to be on the leading edge of what is sure to become a major profession in the not-too-distant future, this book can show you how. Each chapter is filled with practical information that will help you reap the fruits of big data and become a successful Data Scientist. • Learn what big data is and how it differs from traditional data through its main characteristics • Explore the different types of Data Scientists and the skillset each one has. • Dig into what the role of the Data Scientist requires in terms of the relevant mindset, technical skills, experience, and how the Data Scientist connects with other people. • Be a Data Scientist for a day, examining the problems you may encounter and how you tackle them, what programs you use, and how you expand your knowledge and know-how. • See how you can become a Data Scientist, based on

machine learning, or data-related background. • Follow step-by-step through the process of landing a Data Scientist job: where you need to look, how you would present yourself to a potential employer, and what it takes to follow a freelancer path. • Read the case studies of experienced, senior-level Data Scientists, in an attempt to get a better perspective of what this role is, in practice. At the end of the book, there is a glossary of the most important terms that have been introduced.

A Guide for Business Users and Data Scientists

How to Think, Speak, and Understand Data Science, Statistics, and Machine Learning

Hiring Data Scientists and Machine Learning Engineers

Understand, Manage, and Prevent Algorithmic Bias

The Ultimate Step-by-Step Guide to Python Programming. Discover How to Master Big Data Analysis and Understand Machine Learning

Practical Data Science

"Turn yourself into a Data Head. You'll become a more valuable employee and make your organization more successful." Thomas H. Davenport, Research Fellow, Author of Competing on Analytics, Big Data @ Work, and The AI Advantage You've heard the hype around data—now get the facts. In *Becoming a Data Head: How to Think, Speak, and Understand Data Science, Statistics, and Machine Learning*, award-winning data scientists Alex Gutman and Jordan Goldmeier pull back the curtain on data science and give you the language and tools necessary to talk and think critically about it. You'll learn how to: Think statistically and understand the role variation plays in your life and decision making Speak intelligently and ask the right questions about the statistics and results you encounter in the workplace Understand what's really going on with machine learning, text analytics, deep learning, and artificial intelligence Avoid common pitfalls when working with and interpreting data *Becoming a Data Head* is a complete guide for data science in the workplace: covering everything from the personalities you'll work with to the math behind the algorithms. The authors have spent years in data trenches and sought to create a fun, approachable, and eminently readable book. Anyone can become a Data Head—an active participant in data science, statistics, and machine learning. Whether you're a business professional, engineer, executive, or aspiring data scientist, this book is for you.

Start your Data Science career using Python today! Are you ready to start your new exciting career? Ready to crush your machine learning career goals? Are you overwhelmed with complexity of the books on this subject?Then let this breezy and fun little book on Python and machine learning models make you a data scientist in 7 days! First part of this book introduces Python basics including: 1) Data Structures like Pandas 2) Foundational libraries like Numpy, Seaborn and Scikit-Learn Second part of this book shows you how to build predictive machine learning models step by step using techniques such as: 1) Regression analysis 2) Decision tree analysis 3) Training and testing data models 4) And much more! After reading this book you will be able to: 1) Code in Python with confidence 2) Build new machine learning models from scratch 3) Know how to clean and prepare your data for analytics 4) Speak confidently about statistical analysis techniques Data Science was ranked the fast-growing field by LinkedIn and Data Scientist is one of the most highly sought after and lucrative careers in the world! If you are on the fence about making the leap to a new and lucrative career, this is the book for you! What sets this book apart from other books on the topic of Python and Machine learning: 1) Step by step code examples and explanation 2) Complex concepts explained visually 3) Real world applicability of the machine learning models introduced 4) Bonus free code samples that you can try yourself without any prior experience in Python! What do I need to get started? You will have a step by step action plan in place once you finish this book and finally feel that you, can master data science and machine learning and start lucrative and rewarding career! Ready to dive in to the exciting world of Python and Machine Learning? Then scroll up to the top and hit that BUY BUTTON!

Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You'll explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and Datasets—Spark's core APIs—through worked examples Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark clusters and applications Learn the power of Structured Streaming, Spark's stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation

Here's the Perfect Solution if You Want to Become the Master of Data Science and Learn Python Step-by-Step Would you like to: Learn a super competitive skill? Become irreplaceable in the future job market? Upgrade yourself to the ultimate data whizz? If so, then keep reading! Data science is one of the emerging technologies that is set to radically transform the job market. With applications in almost every industry, data science experts will have no shortage of great job offers. But, the whole field may seem a little intimidating if your background is not specific to data science. This book is here to guide you through the field of data science from the very beginning. You will learn the fundamental skills and tools to support your learning process. If you're a beginner, this is the book to help you easily understand the basics of data science. To understand data science, you also need a good understanding of how Python helps you design and implement these projects. This guidebook is going to explain how we can get all of this done. Here just a little preview of what you'll find inside this book: A thorough and simple explanation of data science and the way it works Basics of data science and fundamental skills you need to get started Data science libraries you need to learn to become a data whizz A blueprint for the most used frameworks for Python data science How to process and understand the data and design your own projects AND SO MUCH MORE! Even if you're an absolute beginner with little programming experience, you will find this book easy to follow and implement. This guide is your first step towards a successful data science career, so don't hesitate! Scroll Up, Click the "Buy Now with 1-Click", and Get Your Copy!

Data Science in Education Using R

Data Science for Business

A Guide for Data Scientists

Data Science Job: How to become a Data Scientist

Game Data Science

A hands-on guide to machine learning in Azure and passing the Microsoft Certified DP-100 exam

A Data Scientist's Guide to Acquiring, Cleaning, and Managing Data in R

Machine learning has become an integral part of many commercial applications and research projects, but this field is not exclusive to large companies with extensive research teams. If you use Python, even as a beginner, this book will teach you practical ways to build your own machine learning solutions. With all the data available today, machine learning applications are limited only by your imagination. You'll learn the steps necessary to create a successful machine-learning application with Python and the scikit-learn library. Authors Andreas Müller and Sarah Guido focus on the practical aspects of using machine learning algorithms, rather than the math behind them. Familiarity with the NumPy and matplotlib libraries will help you get even more from this book. With this book, you'll learn: Fundamental concepts and applications of machine learning Advantages and shortcomings of widely used machine learning algorithms How to represent data processed by machine learning, including which data aspects to focus on Advanced methods for model evaluation and parameter tuning The concept of pipelines for chaining models and encapsulating your workflow Methods for working with text data, including text-specific processing techniques Suggestions for improving your machine learning and data science skills

An introduction to data visualization and data storytelling. This book explains (visually) the fundamental principles of a meaningful chart making at high level. No coding or statistics skills required. Audience: data visualization students, senior data scientists, prescriptive analytics consultants. Written by a design thinking professor and multiple-times awarded kaggle master, this book hits the sweet spot between abstraction and detail.

Summary Think Like a Data Scientist presents a step-by-step approach to data science, combining analytic, programming, and business perspectives into easy-to-digest techniques and thought processes for solving real world data-centric problems. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Data collected from customers, scientific measurements, IoT sensors, and so on is valuable only if you understand it. Data scientists revel in the interesting and rewarding challenge of observing, exploring, analyzing, and interpreting this data. Getting started with data science means more than mastering analytic tools and techniques, however; the real magic happens when you begin to think like a data scientist. This book will get you there. About the Book Think Like a Data Scientist teaches you a step-by-step approach to solving real-world data-centric problems. By breaking down carefully crafted examples, you'll learn to combine analytic, programming, and business perspectives into a repeatable process for extracting real knowledge from data. As you read, you'll discover (or remember) valuable statistical techniques and explore powerful data science software. More importantly, you'll put this knowledge together using a structured process for data science. When you've finished, you'll have a strong foundation for a lifetime of data science learning and practice. What's Inside The data science process, step-by-step How to anticipate problems Dealing with uncertainty Best practices in software and scientific thinking About the Reader Readers need beginner programming skills and knowledge of basic statistics. About the Author Brian Godsey has worked in software, academia, finance, and defense and has launched several data-centric start-ups. Table of Contents PART 1 - PREPARING AND GATHERING DATA AND KNOWLEDGE Philosophies of data science Setting goals by asking good questions Data all around us: the virtual wilderness Data wrangling: from capture to domestication Data assessment: poking and prodding PART 2 - BUILDING A PRODUCT WITH SOFTWARE AND STATISTICS Developing a plan Statistics and modeling: concepts and foundations Software: statistics in action Supplementary software: bigger, faster, more efficient Plan execution: putting it all together PART 3 - FINISHING OFF THE PRODUCT AND WRAPPING UP Delivering a product After product delivery: problems and revisions Wrapping up: putting the project away

Game data science, defined as the practice of deriving insights from game data, has created a revolution in the multibillion-dollar games industry - informing and enhancing production, design, and development processes. Almost all game companies and academics have now adopted some type of game data science, every tool utilized by game developers allows collecting data from games, yet there has been no definitive resource for academics and professionals in this rapidly developing sector until now. Games Data Science delivers an excellent introduction to this new domain and provides the definitive guide to methods and practices of computer science, analytics, and data science as applied to video games. It is the ideal resource for academic students and professional learners seeking to understand how data science is used within the game development and production cycle, as well as within the interdisciplinary field of games research. Organized into chapters that integrate laboratory and game data examples, this book provides a unique resource to train and educate both industry professionals and academics about the use of game data science, with practical exercises and examples on how such processes are implemented and used in academia and industry, interweaving theoretical learning with practical application throughout.

The Woefully Complete Guide

Practical Statistics for Data Scientists

Data Science For Dummies

Tackle the data science process step-by-step

Python for Data Science

Think Like a Data Scientist

Developing Analytic Talent

Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. Youd' explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and DatasetsSpark's core APIsthrough worked examples Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark clusters and applications Learn the power of Structured Streaming, Spark's stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation

Data science is the application of a combination of mathematical, statistical, analytical and programming skills for the collection, organization, and interpretation of data to allow effective and proper management of the business whose data it is. The job of such a scientist is trending all over the world. The demand for such scientists is huge, more than the number of available candidates. A recent report explained that the need for these scientists has increased by more than 50% since last year. These scientists often referred to as big data wranglers, are a perfect blend of mathematician and computer scientist. Data science is a field of study that is growing at a fast pace. From big tech companies to E-commerce companies to websites and many others are now relying on data science. Amount of data that is collected by these companies are without any bounds. Semi-structure to big unstructured data is stored in large frameworks of these companies. Now the question is how to use this. What you will gain as knowledge in this book: Why Is Data Science Widely Used? Why Should You Study Data Science? Why Should One Consider Data Science As A Career? Data Science: An Exciting Career Option Types of Data Loss and Recovery Options Data Science and Its Wide Range of Applications What Are the Programming Languages Required for Data Science? Meaning of Data Science in Depth 4 Weird Ways How Data Is Used Around the World 5 Reasons Why Data Science Could Be the Advertising Wave of the Future It is a field where one should be trained and practiced. Without proper training and applicative skills, one cannot be as successful as a data scientist. There is a lot to learn about various data science tools and techniques. Getting certified will not only help you hone your skills but also will confirm your future as a data scientist.

The Complete Guide To Data Science For Beginners: Data Science Books You Should Read