

Getting Started With Data Science

Getting Started With Data Science

Downloaded from formapp.investigatiimedia.ro by Luciano & Rice

WELCOME TO FORMAPP.INVESTIGATIIMEDIA.RO BOOK SHOP!

At our book shop, we believe that **Getting Started With Data Science** have the power to carry us to brand-new globes and spark our passions. That's why we provide a **variety** of **books** that deal with all **passions** and **analysis levels**. Whether you're an experienced visitor or just beginning your literary trip, we have something for every person.

Our collection includes both physical and **Getting Started With Data Science digital publications**, along with **audiobooks**, so you can pick the style that ideal suits your choices. We likewise supply **individualized publication referrals** based upon your **passions** and previous analysis experiences.

Join our area of **publication enthusiasts** and **connect** with like-minded people who share your passion for literary works. We take satisfaction in showcasing new skill and **emerging voices** in the literary globe, so you can discover promising authors that are pushing borders and redefining genres.

Shop with us today and find the delight and expertise that **Getting Started With Data Science books** bring.

Secret Takeaways of Getting Started With Data Science

- Our store provides a **wide array** of Getting Started With Data Science in numerous formats, including physical and **digital books** and **audiobooks**.
- We provide **tailored publication suggestions** based on your **passions** and past analysis experiences.
- Join our area of **Getting Started With Data Science book enthusiasts** to **link** with similar people and increase your literary perspectives.
- We showcase new talent and **arising voices** in the literary globe, so you can uncover promising writers who are pressing borders and redefining genres.
- **Discover** our collection and discover the delight and expertise that Getting Started With Data Science publications bring.

EXPLORE OUR WIDE ARRAY OF GETTING STARTED WITH DATA SCIENCE!

At our store, we satisfaction ourselves available a substantial series of publications in numerous categories. Our company believe that everybody should have access to literary works that talks with them and enhances their lives. That's why we make every effort to have something for everybody in our collection.

Whether you're a follower of fiction, non-fiction, love, mystery, or any other category, we have something to please your analysis yearnings. Our shelves are stocked with both timeless and modern titles, and our well-informed personnel is always handy to aid assist you in the right instructions.

Our team believe that reading is a journey of expedition, and we intend to help you start that journey in the best feasible way. That's why we offer a **wide array** of Getting Started With Data Science to choose from, so you can discover the ideal suit for your passions and reading choices.

Discover our variety of this Getting Started With Data Science books!

- Discover a large range of genres
- Discover the **excellent book** for your interests
- Dive into **captivating storytelling of Getting Started With Data Science**
- **Discover brand-new authors** and **emerging voices**
- Find books for **every ages** and **analysis levels**

So whether you're looking for an awesome page-turner or a thought-provoking literary work of art, we have something for everybody. Come **check out** our wide range of publications and let us assist you find your next fantastic read!

DIVE INTO CAPTIVATING NARRATION

At our shop, our company believe that Getting Started With Data Science are more than simply ink on paper - they're a portal to new worlds and experiences. That's why we meticulously curate our selection to consist of publications with **exciting storytelling** that will certainly transport you to various places and times.

From sprawling legends to heartfelt love, our publications are full of well-developed characters and thought-provoking styles that will certainly keep you engaged from beginning to end. Whether you're seeking a retreat from fact or a deeper understanding of the world around you, our Getting Started With Data Science publications offer something for everyone.

Our writers are masters of their craft, making use of language and images to produce vivid worlds that feel genuine. With every turn of the page, you'll be drawn deeper into the story, anxious to discover what happens next.

Experience the power of fascinating storytelling of Getting Started With Data Science

Our company believe that analysis is a transformative experience that can alter your life. By diving into a **Getting Started With Data Science book** with **fascinating storytelling**, you have the power to discover brand-new points, acquire brand-new viewpoints, and **link** with personalities and circumstances that might be various from your own.

Whether you're a seasoned reader or simply getting started, our choice of books with captivating storytelling makes sure to leave an enduring perception. We welcome you to **discover** our store and find the magic of narration for yourself.

DISCOVER NEW AUTHORS AND EMERGING VOICES

At our shop, we take satisfaction in showcasing new talent and emerging voices in the literary globe. We believe that every terrific tale should have to be informed which the power of literary works lies in its ability to magnify varied viewpoints and voices.

By discovering our choice, you'll have the opportunity to find and sustain encouraging writers who are pushing limits and redefining categories. From thought-provoking memoirs to exhilarating debut stories, our collection spans a wide range of genres and styles, ensuring that there's something for everybody.

So why not step outdoors your reading convenience zone and discover a new author or emerging voice today? You never ever understand what literary treasures you may discover!

TAKE PLEASURE IN THE BENEFIT OF PHYSICAL AND GETTING STARTED WITH DATA SCIENCE DIGITAL BOOKS

At our shop, we comprehend that every visitor has their very own preferences when it pertains to the style they choose to review their publications. That's why we offer both physical and **digital publications of Getting Started With Data Science** for your **ease**.

The Benefit of Physical Books

For those who like the feeling of a **Getting Started With Data Science book** in their hands and the experience of scanning the web pages, we provide a wide variety of **physical books**. From hardbounds to paperbacks, our selection consists of all styles and **reading degrees**. You can search our shelves and take your time choosing the ideal publication to contribute to your collection.

The Comfort of Digital Books

We likewise understand that some viewers like the **comfort** of continuing reading a tablet computer or e-reader. That's why we provide a substantial choice of electronic publications that you can purchase and download and install easily. You can take your library with you anywhere you go and delight in the advantages of having your preferred books within your reaches.

Whichever layout you prefer, we have actually got you covered. Our goal is to supply you with the most effective analysis experience while additionally keeping in mind your **comfort**. Shop with us today and uncover the **delight of reading Getting Started With Data Science** in your favored style.

IMPROVE YOUR REVIEWING WITH AUDIOBOOKS

At our shop, our team believe that analysis should be a practical and pleasurable experience for every person. That's why we offer a broad selection of **audiobooks** that you can pay attention to on-the-go or while loosening up in your home. With our audiobook collection, you can enhance your reading experience and submerge on your own in captivating narratives.

Our audiobooks are meticulously curated to give the very same degree of entertainment and knowledge as our physical and electronic publications. Whether you're listening to an exhilarating mystery or a provocative memoir, our audiobooks will certainly move you to brand-new globes and offer

you a fresh point of view.

The Benefits of Audiobooks

Audiobooks provide an unique means to enjoy Getting Started With Data Science and enhance your analysis experience. Here are a few of the advantages of audiobooks:

- Useful for people who have trouble checking out as a result of visual impairment or finding out impairments.
- Great for multitasking while driving, working out, or doing household chores.
- Can boost enunciation, vocabulary, and language understanding skills.
- Practical for people that fight with holding a book or analysis for an extended period of time.

Our audiobook collection includes a selection of genres and authors, so you can discover something that lines up with your reading interests. Whether you're a fan of romance, science fiction, historical fiction, or self-help books, we have options that satisfy every preference.

Exactly how to Gain access to Getting Started With Data Science Audiobook Collection

You can access our audiobook collection through our site or by visiting our physical store. We provide both physical and digital audiobooks to suit your choices. If you select to buy Getting Started With Data Science electronic audiobooks, you can download them quickly and start listening today. Conversely, if you favor physical audiobooks, you can order them online and we'll supply them to your front door.

Improve your reading experience with our audiobook collection today and discover the happiness of paying attention to fascinating narrations that will carry you to brand-new globes.

LOCATE BOOKS FOR ANY AGES AND ANALYSIS LEVELS

As a viewers, you know that publications can be enjoyed by people of **any ages** and reading levels. That is why our collection includes books that accommodate individuals of different ages and reading skills.

For the kids, we have a substantial range of children's publications, consisting of photo books, storybooks, and task books. Our option of young adult novels makes up the most up to date patterns and classic standards that young people can delight in.

If you like publications for fully grown viewers, we have a wide series of literary works that makes certain to pique your rate of interest. You can choose from modern jobs or delve into traditional stories from prominent authors.

Whether you are an experienced bibliophile or just starting on your reading journey, we have a book for each preference and ability degree. Our team of well-informed staff can help you find a Getting Started With Data Science publication that matches your passions and checking out efficiency.

Discover the Perfect Publication for Your Checking Out Level

At our shop, we take satisfaction in accommodating readers of **every ages** and ability levels. We comprehend that selecting the appropriate publication can be an overwhelming job, particularly for beginners. That is why we offer customized suggestions based upon your analysis level and passions.

We want you to feel great in your book selections, and our team of professionals is here to make that happen. From kids's publications to adult literature, you can rely on that we will help you find the ideal publication that matches your analysis skills and passions.

Individualized Recommendations

Our team is here to aid you discover your next literary experience. Get in touch with us and get personalized suggestions based on your passions and checking out choices. We'll deal with you to discover the perfect publication that fits your tastes and leaves you desiring more.

Join our neighborhood of Getting Started With Data Science publication enthusiasts today and get in touch with like-minded people that share your passion for literature. We can't wait to start new literary trips with you!

OBTAIN CUSTOMIZED BOOK SUGGESTIONS

At our store, we understand that picking the ideal book can be frustrating. That's why we offer individualized book referrals based on your interests and reading choices. Our well-informed personnel is always ready to aid and provide customized recommendations to assist you find your following fantastic read.

Whether you're looking for a particular style, author, or design, we'll aid you limit your options and discover a publication that fits your preferences. We put in the time to be familiar with our clients and their analysis behaviors, guaranteeing that each referral is distinctly tailored to their demands.

With our personalized publication suggestions, you'll never ever need to squander your time or money on a publication that does not resonate with you. Let us help you discover your following literary experience and enhance your reading experience.

GIFT THE HAPPINESS OF CHECKING OUT GETTING STARTED WITH DATA SCIENCE

Books make terrific gifts for any type of occasion. At our store, we provide a broad selection of publications for any ages and interests, guaranteeing that you'll find the excellent book for your enjoyed ones.

Absolutely nothing defeats the sensation of providing someone the **happiness of reviewing Getting Started With Data Science**. Whether it's an awesome secret, a heartwarming love, or a fascinating biography, publications have the power to move visitors to brand-new worlds and spark their creative imaginations.

With our personalized referral solution, we can help you select the best book for your recipient. Our knowledgeable personnel will take into account their passions and reviewing choices to offer customized pointers that make certain to pleasure.

Not sure which book to choose? Think about among our **gift** cards, which enables the recipient to explore our collection and select a publication that speaks to them personally.

So why not share the magic of reading with those you enjoy? Give the **Getting Started With Data Science** of a book and make their day a little more vibrant!

STAY UPGRADED WITH THE CURRENT RELEASES AND BESTSELLERS

At our shop, we satisfaction ourselves on maintaining our collection up-to-date with the **Getting Started With Data Science newest releases** and **bestsellers**. We know exactly how essential it is to stay in the loop with the literary globe and have the latest titles at your fingertips.

Whether you're a fan of enigma novels, self-help guides, or romance tales, we have the **most recent launches** and **bestsellers** in every genre. Our group stays on top of brand-new launches and preferred titles, so you can always discover something to read that's fresh and exciting.

Looking for something certain? Our educated team can help you find hard-to-find titles and suggest comparable writers and titles you might enjoy. Whether you're a skilled reader or new to the literary globe, we have something for everybody.

What's New

Several of our recent enhancements to the collection include:

These titles are promptly ending up being brand-new follower favorites and make certain to astound you with their gripping narration and distinct perspectives.

Don't lose out on the **Getting Started With Data Science most current launches** and bestsellers - shop with us today and discover your following favorite read!

DELIGHT IN EXCLUSIVE DISCOUNTS AND PROMOTIONS

At our shop, we believe that every person should have access to fantastic publications at cost effective prices. That's why we offer **Getting Started With Data Science special discount rates** and **promos** to our valued clients.

Enroll in our newsletter to receive updates on the most recent **promotions** and price cuts offered. You can additionally follow us on social media, where we frequently announce special deals and limited-time deals.

As a part of our commitment program, you'll receive much more **unique discounts** and incentives, consisting of totally free shipping and unique grant acquisition.

Whether you're a new client or a dedicated buyer, we want to see to it that you have accessibility to the best possible bargains on our publications. Shop with us today and take advantage of our **unique discounts** and **promos**!

PRACTICAL ONLINE BUYING AND FAST DELIVERY

Purchasing your favorite publications has never been easier! At our store, we provide a **hassle-free on-line buying** system that allows you to search our considerable collection from the comfort of your home. Simply include your preferred books to your cart, enter your shipping information, and proceed to check out. It's that easy!

However the comfort does not stop there. We likewise supply **fast delivery** options that ensure you obtain your books in a prompt manner. Whether you require them for a book club conference or just can not wait to dive into a brand-new tale, we've obtained you covered.

Delivering Alternatives

- Requirement shipment: Our basic distribution alternative generally takes 3-7 organization days to arrive.
- Expedited distribution: Need your publications earlier? Our expedited delivery alternative assurances delivery within 1-3 service days.
- Curbside pickup: If you favor to get your books personally, we also supply curbside pickup at choose areas.

Our team works hard to guarantee your Getting Started With Data Science orders are processed and shipped as swiftly as feasible. Plus, with our easy-to-use online radar, you can watch on your bundle's progress every action of the way.

Experience the ease and speed of our on-line buying and distribution system. Position your order today and take the initial step on your next literary experience!

CONSUMER CONTENTMENT GUARANTEED

At our shop, we take wonderful satisfaction in making sure that our consumers are satisfied with their purchasing experience. From the minute you begin surfing our web site to the final distribution of your order, we are committed to supplying outstanding solution.

If you have any type of questions or concerns about your order, please do not be reluctant to connect to our customer service group. We are right here to aid you and ensure that you are entirely satisfied with your acquisition.

We offer a problem-free return policy for any products that do not satisfy your expectations. We also provide a protected settlement entrance, so you can shop with confidence, recognizing that your individual details is shielded.

We support the quality of Getting Started With Data Science publications and are confident that you will certainly be satisfied with your purchase. If you are not totally satisfied with your order, we will certainly do everything we can to make it right.

Our commitment to client contentment is steady, and we strive to surpass your expectations in every method feasible. Shop with us today and experience the distinction!

CONCLUSION

We wish you're as delighted regarding the world of Getting Started With Data Science books as we are! Our shop is devoted to providing an outstanding reading experience for all. With our wide variety of publications, from brand-new releases to ageless standards, there is something for every person. Our dedication to consumer contentment suggests that you can shop confidently, recognizing that your order is in great hands.

Do not neglect to connect with us and fellow book lovers in our vivid area. Share your favorite reads, take part in thoughtful conversations, and uncover new literary treasures. And if you ever need help discovering the excellent book, our knowledgeable team is below to give customized suggestions.

Getting Started with Data Science

Making Sense of Data with Analytics IBM Press

Master Data Analytics Hands-On by Solving Fascinating Problems You'll Actually Enjoy! Harvard Business Review recently called data science "The Sexiest Job of the 21st Century." It's not just sexy: For millions of managers, analysts, and students who need to solve real business problems, it's indispensable. Unfortunately, there's been nothing easy about learning data science-until now. Getting Started with Data Science takes its inspiration from worldwide best-sellers like Freakonomics and Malcolm Gladwell's Outliers: It teaches through a powerful narrative packed with unforgettable stories. Murtaza Haider offers informative, jargon-free coverage of basic theory and technique, backed with plenty of vivid examples and hands-on practice opportunities. Everything's software and platform agnostic, so you can learn data science whether you work with R, Stata, SPSS, or SAS. Best of all, Haider teaches a crucial skillset most data science books ignore: how to tell powerful stories using graphics and tables. Every chapter is built around real research challenges, so you'll always know why you're doing what you're doing. You'll master data science by answering fascinating questions, such as: • Are religious individuals more or less likely to have extramarital affairs? • Do attractive professors get better teaching evaluations? • Does the higher price of cigarettes deter smoking? • What determines housing prices more: lot size or the number of bedrooms? • How do teenagers and older people differ in the way they use social media? • Who is more likely to use online dating services? • Why do some purchase iPhones and others Blackberry devices? • Does the presence of children influence a family's spending on alcohol? For each problem, you'll walk through defining your question and the answers you'll need; exploring how others have approached similar challenges; selecting your data and methods; generating your statistics; organizing your report; and telling your story. Throughout, the focus is squarely on what matters most: transforming data into insights that are clear, accurate, and can be acted upon.

Getting Started with Data Science

Making Sense of Data with Analytics IBM Press

Master Data Analytics Hands-On by Solving Fascinating Problems You'll Actually Enjoy! Harvard Business Review recently called data science "The Sexiest Job of the 21st Century." It's not just sexy: For millions of managers, analysts, and students who need to solve real business problems, it's indispensable. Unfortunately, there's been nothing easy about learning data science-until now. Getting Started with Data Science takes its inspiration from worldwide best-sellers like Freakonomics and Malcolm Gladwell's Outliers: It teaches through a powerful narrative packed with unforgettable stories. Murtaza Haider offers informative, jargon-free coverage of basic theory and technique, backed with plenty of vivid examples and hands-on practice opportunities. Everything's software and platform agnostic, so you can learn data science whether you work with R, Stata, SPSS, or SAS. Best of all, Haider teaches a crucial skillset most data science books ignore: how to tell powerful stories using graphics and tables. Every chapter is built around real research challenges, so you'll always know why you're doing what you're doing. You'll master data science by answering fascinating

questions, such as: * Are religious individuals more or less likely to have extramarital affairs? * Do attractive professors get better teaching evaluations? * Does the higher price of cigarettes deter smoking? * What determines housing prices more: lot size or the number of bedrooms? * How do teenagers and older people differ in the way they use social media? * Who is more likely to use online dating services? * Why do some purchase iPhones and others Blackberry devices? * Does the presence of children influence a family's spending on alcohol? For each problem, you'll walk through defining your question and the answers you'll need; exploring how others have approached similar challenges; selecting your data and methods; generating your statistics; organizing your report; and telling your story. Throughout, the focus is squarely on what matters most: transforming data into insights that are clear, accurate, and can be acted upon.

Getting Started with Data Science

Making Sense of Data with Analytics

Getting Started with Streamlit for Data Science

Create and deploy Streamlit web applications from scratch in Python Packt Publishing Ltd

Create, deploy, and test your Python applications, analyses, and models with ease using Streamlit Key Features Learn how to showcase machine learning models in a Streamlit application effectively and efficiently Become an expert Streamlit creator by getting hands-on with complex application creation Discover how Streamlit enables you to create and deploy apps effortlessly Book Description Streamlit shortens the development time for the creation of data-focused web applications, allowing data scientists to create web app prototypes using Python in hours instead of days. Getting Started with Streamlit for Data Science takes a hands-on approach to helping you learn the tips and tricks that will have you up and running with Streamlit in no time. You'll start with the fundamentals of Streamlit by creating a basic app and gradually build on the foundation by producing high-quality graphics with data visualization and testing machine learning models. As you advance through the chapters, you'll walk through practical examples of both personal data projects and work-related data-focused web applications, and get to grips with more challenging topics such as using Streamlit Components, beautifying your apps, and quick deployment of your new apps. By the end of this book, you'll be able to create dynamic web apps in Streamlit quickly and effortlessly using the power of Python. What you will learn Set up your first development environment and create a basic Streamlit app from scratch Explore methods for uploading, downloading, and manipulating data in Streamlit apps Create dynamic visualizations in Streamlit using built-in and imported Python libraries Discover strategies for creating and deploying machine learning models in Streamlit Use Streamlit sharing for one-click deployment Beautify Streamlit apps using themes, Streamlit Components, and Streamlit sidebar Implement best practices for prototyping your data science work with Streamlit Who this book is for This book is for data scientists and machine learning enthusiasts who want to create web apps using Streamlit. Whether you're a junior data scientist looking to deploy your first machine learning project in Python to improve your resume or a senior data scientist who wants to use Streamlit to make convincing and dynamic data analyses, this book will help you get there! Prior knowledge of Python programming will assist with understanding the concepts covered.

R for Data Science

Import, Tidy, Transform, Visualize, and Model Data "O'Reilly Media, Inc."

"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"--

Data Science in Education Using R Routledge

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

Build a Career in Data Science Simon and Schuster

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 search: Identifying the right job for you
6. The application: Résumés and cover letters
7. The interview: What to expect and how to handle it
8. The offer: Knowing what 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

Data Science For Dummies *John Wiley & Sons*

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

Data Science from Scratch

First Principles with Python *"O'Reilly Media, Inc."*

Data science libraries, frameworks, modules, and toolkits are great for doing data science, but they’re also a good way to dive into the discipline without actually understanding data science. In this book, you'll learn how many of the most fundamental data science tools and algorithms work by implementing them from scratch. If you have an aptitude for mathematics and some programming skills, author Joel Grus will help you get comfortable with the math and statistics at the core of data science, and with hacking skills you need to get started as a data scientist. Today’s messy glut of data holds answers to questions no one’s even thought to ask. This book provides you with the know-how to dig those answers out. Get a crash course in Python Learn the basics of linear algebra, statistics, and probability—and understand how and when they're used in data science Collect, explore, clean, munge, and manipulate data Dive into the fundamentals of machine learning Implement models such as k-nearest Neighbors, Naive Bayes, linear and logistic regression, decision trees, neural networks, and clustering Explore recommender systems, natural language processing, network analysis, MapReduce, and databases

Statistical Inference via Data Science: A ModernDive into R and the Tidyverse *CRC Press*

Statistical Inference via Data Science: A ModernDive into R and the Tidyverse provides a pathway for learning about statistical inference using data science tools widely used in industry, academia, and government. It introduces the tidyverse suite of R packages, including the ggplot2 package for data visualization, and the dplyr package for data wrangling. After equipping readers with just enough of these data science tools to perform effective exploratory data analyses, the book covers traditional introductory statistics topics like confidence intervals, hypothesis testing, and multiple regression modeling, while focusing on visualization throughout. Features: ● Assumes minimal prerequisites, notably, no prior calculus nor coding experience ● Motivates theory using real-world data, including all domestic flights leaving New York City in 2013, the Gapminder project, and the data journalism website, FiveThirtyEight.com ● Centers on simulation-based approaches to statistical inference rather than mathematical formulas ● Uses the infer package for "tidy" and transparent statistical inference to construct confidence intervals and conduct hypothesis tests via the bootstrap and permutation methods ● Provides all code and output embedded directly in the text; also available in the online version at moderndive.com This book is intended for individuals who would like to simultaneously start developing their data science toolbox and start learning about the inferential and modeling tools used in much of modern-day research. The book can be used in methods and data science courses and first courses in statistics, at both the undergraduate and graduate levels.

Introduction to Data Science

Data Analysis and Prediction Algorithms with R *CRC Press*

Introduction to Data Science: Data Analysis and Prediction Algorithms with R introduces concepts and skills that can help you tackle real-world data analysis challenges. It covers concepts from probability, statistical inference, linear regression, and machine learning. It also helps you develop skills such as R programming, data wrangling, data visualization, predictive algorithm building, file organization with UNIX/Linux shell, version control with Git and GitHub, and reproducible document preparation. This book is a textbook for a first course in data science. No previous knowledge of R is necessary, although some experience with programming may be helpful. The book is divided into six parts: R, data visualization, statistics with R,

data wrangling, machine learning, and productivity tools. Each part has several chapters meant to be presented as one lecture. The author uses motivating case studies that realistically mimic a data scientist’s experience. He starts by asking specific questions and answers these through data analysis so concepts are learned as a means to answering the questions. Examples of the case studies included are: US murder rates by state, self-reported student heights, trends in world health and economics, the impact of vaccines on infectious disease rates, the financial crisis of 2007-2008, election forecasting, building a baseball team, image processing of hand-written digits, and movie recommendation systems. The statistical concepts used to answer the case study questions are only briefly introduced, so complementing with a probability and statistics textbook is highly recommended for in-depth understanding of these concepts. If you read and understand the chapters and complete the exercises, you will be prepared to learn the more advanced concepts and skills needed to become an expert.

Python for Data Analysis

Data Wrangling with Pandas, NumPy, and IPython *"O'Reilly Media, Inc."*

Get complete instructions for manipulating, processing, cleaning, and crunching datasets in Python. Updated for Python 3.6, the second edition of this hands-on guide is packed with practical case studies that show you how to solve a broad set of data analysis problems effectively. You’ll learn the latest versions of pandas, NumPy, IPython, and Jupyter in the process. Written by Wes McKinney, the creator of the Python pandas project, this book is a practical, modern introduction to data science tools in Python. It’s ideal for analysts new to Python and for Python programmers new to data science and scientific computing. Data files and related material are available on GitHub. Use the IPython shell and Jupyter notebook for exploratory computing Learn basic and advanced features in NumPy (Numerical Python) Get started with data analysis tools in the pandas library Use flexible tools to load, clean, transform, merge, and reshape data Create informative visualizations with matplotlib Apply the pandas groupby facility to slice, dice, and summarize datasets Analyze and manipulate regular and irregular time series data Learn how to solve real-world data analysis problems with thorough, detailed examples

Think Like a Data Scientist

Tackle the data science process step-by-step *Simon and Schuster*

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

The Data Science Workshop

A New, Interactive Approach to Learning Data Science *Packt Publishing Ltd*

Cut through the noise and get real results with a step-by-step approach to data science Key Features Ideal for the data science beginner who is getting started for the first time A data science tutorial with step-by-step exercises and activities that help build key skills Structured to let you progress at your own pace, on your own terms Use your physical print copy to redeem free access to the online interactive edition Book Description You already know you want to learn data science, and a smarter way to learn data science is to learn by doing. The Data Science Workshop focuses on building up your practical skills so that you can understand how to develop simple machine learning models in Python or even build an advanced model for detecting potential bank frauds with effective modern data science. You'll learn from real examples that lead to real results. Throughout The Data Science Workshop, you'll take an engaging step-by-step approach to understanding data science. You won't have to sit through any unnecessary theory. If you're short on time you can jump into a single exercise each day or spend an entire weekend training a model using sci-kit learn. It's your choice. Learning on your terms, you'll build up and reinforce key skills in a way that feels rewarding. Every physical print copy of The Data Science Workshop unlocks access to the interactive edition. With videos detailing all exercises and activities, you'll always have a guided solution. You can also benchmark yourself against assessments, track progress, and receive content updates. You'll even earn a secure credential

that you can share and verify online upon completion. It's a premium learning experience that's included with your printed copy. To redeem, follow the instructions located at the start of your data science book. Fast-paced and direct, The Data Science Workshop is the ideal companion for data science beginners. You'll learn about machine learning algorithms like a data scientist, learning along the way. This process means that you'll find that your new skills stick, embedded as best practice. A solid foundation for the years ahead. What you will learn Find out the key differences between supervised and unsupervised learning Manipulate and analyze data using scikit-learn and pandas libraries Learn about different algorithms such as regression, classification, and clustering Discover advanced techniques to improve model ensembling and accuracy Speed up the process of creating new features with automated feature tool Simplify machine learning using open source Python packages Who this book is for Our goal at Packt is to help you be successful, in whatever it is you choose to do. The Data Science Workshop is an ideal data science tutorial for the data science beginner who is just getting started. Pick up a Workshop today and let Packt help you develop skills that stick with you for life.

Python Data Science Handbook

Essential Tools for Working with Data *"O'Reilly Media, Inc."*
For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the Python Data Science Handbook do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you'll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine learning algorithms

Data Science at the Command Line

Facing the Future with Time-Tested Tools *"O'Reilly Media, Inc."*
This hands-on guide demonstrates how the flexibility of the command line can help you become a more efficient and productive data scientist. You'll learn how to combine small, yet powerful, command-line tools to quickly obtain, scrub, explore, and model your data. To get you started—whether you're on Windows, OS X, or Linux—author Jeroen Janssens introduces the Data Science Toolbox, an easy-to-install virtual environment packed with over 80 command-line tools. Discover why the command line is an agile, scalable, and extensible technology. Even if you're already comfortable processing data with, say, Python or R, you'll greatly improve your data science workflow by also leveraging the power of the command line. Obtain data from websites, APIs, databases, and spreadsheets Perform scrub operations on plain text, CSV, HTML/XML, and JSON Explore data, compute descriptive statistics, and create visualizations Manage your data science workflow using Drake Create reusable tools from one-liners and existing Python or R code Parallelize and distribute data-intensive pipelines using GNU Parallel Model data with dimensionality reduction, clustering, regression, and classification algorithms

The Big R-Book

From Data Science to Learning Machines and Big Data *John Wiley & Sons*
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.

Data Science and Big Data Analytics

Discovering, Analyzing, Visualizing and Presenting Data *John Wiley & Sons*
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!

Getting Started with Business Analytics

Insightful Decision-Making *CRC Press*
Assuming no prior knowledge or technical skills, Getting Started with Business Analytics: Insightful Decision-Making explores the contents, capabilities, and applications of business analytics. It bridges the worlds of business and statistics and describes business analytics from a non-commercial standpoint. The authors demystify the main concepts and terminologies and give many examples of real-world applications. The first part of the book introduces business data and recent technologies that have promoted fact-based decision-making. The authors look at how business intelligence differs from business analytics. They also discuss the main components of a business analytics application and the various requirements for integrating business with analytics. The second part presents the technologies underlying business analytics: data mining and data analytics. The book helps you understand the key concepts and ideas behind data mining and shows how data mining has expanded into data analytics when considering new types of data such as network and text data. The third part explores business analytics in depth, covering customer, social, and operational analytics. Each chapter in this part incorporates hands-on projects based on publicly available data. Helping you make sound decisions based on hard data, this self-contained guide provides an integrated framework for data mining in business analytics. It takes you on a journey through this data-rich world, showing you how to deploy business analytics solutions in your organization.

Python for Data Science For Dummies *John Wiley & Sons*
The fast and easy way to learn Python programming and statistics Python is a general-purpose programming language created in the late 1980s—and named after Monty Python—that's used by thousands of people to do things from testing microchips at Intel, to powering Instagram, to building video games with the PyGame library. Python For Data Science For Dummies is written for people who are new to data analysis, and discusses the basics of Python data analysis programming and statistics. The book also discusses Google Colab, which makes it possible to write Python code in the cloud. Get started with data science and Python Visualize information Wrangle data Learn from data The book provides the statistical background needed to get started in data science programming, including probability, random distributions, hypothesis testing, confidence intervals, and building regression models for prediction.

Programming Skills for Data Science

Start Writing Code to Wrangle, Analyze, and Visualize Data with R *Addison-Wesley Professional*
The Foundational Hands-On Skills You Need to Dive into Data Science “Freeman and Ross have created the definitive resource for new and aspiring data scientists to learn foundational programming skills.” –From the foreword by Jared Lander, series editor Using data science techniques, you can transform raw data into actionable insights for domains ranging from urban planning to precision medicine. Programming Skills for Data Science brings together all the foundational skills you need to get started, even if you have no programming or data science experience. Leading instructors Michael Freeman and Joel Ross guide you through installing and configuring the tools you need to solve professional-level data science problems, including the widely used R language and Git version-control system. They explain how to wrangle your data into a form where it can be easily used, analyzed, and visualized so others can see the patterns you’ve uncovered. Step by step, you’ll master powerful R programming techniques and troubleshooting skills for probing data in new ways, and at larger scales. Freeman and Ross teach through practical examples and exercises that can be combined into complete data science projects. Everything’s focused on real-world application, so you can quickly start analyzing your own data and getting answers you can act upon. Learn to Install your complete data science environment, including R and RStudio Manage projects efficiently, from version tracking to documentation Host, manage, and collaborate on data science projects with GitHub Master R language fundamentals: syntax, programming concepts, and data structures Load, format, explore, and restructure data for successful analysis Interact with databases and web APIs Master key principles for visualizing data accurately and intuitively Produce engaging, interactive visualizations with ggplot and other R packages Transform analyses into sharable documents and sites with R Markdown Create interactive web data science applications with Shiny Collaborate smoothly as part of a data science team Register your book for convenient access to downloads, updates, and/or corrections as they become available. See inside book for details.

Principles of Data Science *Packt Publishing Ltd*
Learn the techniques and math you need to start making sense of your data About This Book Enhance your knowledge of coding with data science theory for practical insight into data science and analysis More than just a math class, learn how to perform real-world data science tasks with R and Python Create actionable insights and transform raw data into tangible value Who This Book Is For You should be fairly well acquainted with basic algebra and should feel comfortable reading snippets of R/Python as well as pseudo code. You should have the urge to learn and apply the techniques

put forth in this book on either your own data sets or those provided to you. If you have the basic math skills but want to apply them in data science or you have good programming skills but lack math, then this book is for you. What You Will Learn Get to know the five most important steps of data science Use your data intelligently and learn how to handle it with care Bridge the gap between mathematics and programming Learn about probability, calculus, and how to use statistical models to control and clean your data and drive actionable results Build and evaluate baseline machine learning models Explore the most effective metrics to determine the success of your machine learning models Create data visualizations that communicate actionable insights Read and apply machine learning concepts to your problems and make actual predictions In Detail Need to turn your skills at programming into effective data science skills? Principles of Data Science is created to help you join the dots between mathematics, programming, and business analysis. With this book, you'll feel confident about asking—and answering—complex and sophisticated questions of your data to move from abstract and raw statistics to actionable ideas. With a unique approach that bridges the gap between mathematics and computer science, this books takes you through the entire data science pipeline. Beginning with cleaning and preparing data, and effective data mining strategies and techniques, you'll move on to build a comprehensive picture of how every piece of the data science puzzle fits together. Learn the fundamentals of computational mathematics and statistics, as well as some pseudocode being used today by data scientists and analysts. You'll get to grips with machine learning, discover the statistical models that help you take control and navigate even the densest datasets, and find out how to create powerful visualizations that communicate what your data means. Style and approach This is an easy-to-understand and accessible tutorial. It is a step-by-step guide with use cases, examples, and illustrations to get you well-versed with the concepts of data science. Along with explaining the fundamentals, the book will also introduce you to slightly advanced concepts later on and will help you implement these techniques in the real world.

Python Machine Learning *Packt Publishing Ltd*
Unlock deeper insights into Machine Leaning with this vital guide to cutting-edge predictive analytics About This Book Leverage Python's most powerful open-source libraries for deep learning, data wrangling, and data visualization Learn effective strategies and best practices to improve and optimize machine learning systems and algorithms Ask – and answer – tough questions of your data with robust statistical models, built for a range of datasets Who This Book Is For If you want to find out how to use Python to start answering critical questions of your data, pick up Python Machine Learning – whether you want to get started from scratch or want to extend your data science knowledge, this is an essential and unmissable resource. What You Will Learn Explore how to use different machine learning models to ask different questions of your data Learn how to build neural networks using Keras and Theano Find out how to write clean and elegant Python code that will optimize the strength of your algorithms Discover how to embed your machine learning model in a web application for increased accessibility Predict continuous target outcomes using regression analysis Uncover hidden patterns and structures in data with clustering Organize data using effective pre-processing techniques Get to grips with sentiment analysis to delve deeper into textual and social media data In Detail Machine learning and predictive analytics are transforming the way businesses and other organizations operate. Being able to understand trends and patterns in complex data is critical to success, becoming one of the key strategies for unlocking growth in a challenging contemporary marketplace. Python can help you deliver key insights into your data – its unique capabilities as a language let you build sophisticated algorithms and statistical models that can reveal new perspectives and answer key questions that are vital for success. Python Machine Learning gives you access to the world of predictive analytics and demonstrates why Python is one of the world's leading data science languages. If you want to ask better questions of data, or need to improve and extend the capabilities of your machine learning systems, this practical data science book is invaluable. Covering a wide range of powerful Python libraries, including scikit-learn, Theano, and Keras, and featuring guidance and tips on everything from sentiment analysis to neural networks, you'll soon be able to answer some of the most important questions facing you and your organization. Style and approach Python Machine Learning connects the fundamental theoretical principles behind machine learning to their practical application in a way that focuses you on asking and answering the right questions. It walks you through the key elements of Python and its powerful machine learning libraries, while demonstrating how to get to grips with a range of statistical models.

Data Science on AWS "*O'Reilly Media, Inc.*"
With this practical book, AI and machine learning practitioners will learn how to successfully build and deploy data science projects on Amazon Web Services. The Amazon AI and machine learning stack unifies data science, data engineering, and application development to help level upyour skills. This guide shows you how to build and run pipelines in the cloud, then integrate the results into applications in minutes instead of days. Throughout the book, authors Chris Fregly and Antje Barth demonstrate how to reduce cost and improve performance. Apply the Amazon AI and ML stack to real-world use cases for natural language processing, computer vision, fraud detection, conversational devices, and more Use automated machine learning to implement a specific subset of use cases with SageMaker Autopilot Dive deep into the complete model development lifecycle for a BERT-based NLP use case including data ingestion, analysis, model training, and deployment Tie everything together into a repeatable machine learning operations pipeline Explore real-time ML, anomaly detection, and streaming analytics on data streams with Amazon Kinesis and Managed Streaming for Apache Kafka Learn security best practices for data science projects and workflows including identity and access management, authentication, authorization, and more

Data Science at the Command Line "*O'Reilly Media, Inc.*"
This thoroughly revised guide demonstrates how the flexibility of the command line can help you become a more efficient and productive data scientist. You'll learn how to combine small yet powerful command-line tools to quickly obtain, scrub, explore, and model your data. To get you started, author Jeroen Janssens provides a Docker image packed with over 80 tools--useful whether you work with Windows, macOS, or Linux. You'll quickly discover why the command line is an agile, scalable, and extensible technology. Even if you're comfortable processing data with Python or R, you'll learn how to greatly improve your data science workflow by leveraging the command line's power. This book is ideal for data scientists, analysts, and engineers; software and machine learning engineers; and system administrators. Obtain data from websites, APIs, databases, and

spreadsheets Perform scrub operations on text, CSV, HTM, XML, and JSON files Explore data, compute descriptive statistics, and create visualizations Manage your data science workflow Create reusable command-line tools from one-liners and existing Python or R code Parallelize and distribute data-intensive pipelines Model data with dimensionality reduction, clustering, regression, and classification algorithms

Data Science and Machine Learning

Mathematical and Statistical Methods *CRC Press*
"This textbook is a well-rounded, rigorous, and informative work presenting the mathematics behind modern machine learning techniques. It hits all the right notes: the choice of topics is up-to-date and perfect for a course on data science for mathematics students at the advanced undergraduate or early graduate level. This book fills a sorely-needed gap in the existing literature by not sacrificing depth for breadth, presenting proofs of major theorems and subsequent derivations, as well as providing a copious amount of Python code. I only wish a book like this had been around when I first began my journey!" -Nicholas Hoell, University of Toronto "This is a well-written book that provides a deeper dive into data-scientific methods than many introductory texts. The writing is clear, and the text logically builds up regularization, classification, and decision trees. Compared to its probable competitors, it carves out a unique niche. -Adam Loy, Carleton College The purpose of Data Science and Machine Learning: Mathematical and Statistical Methods is to provide an accessible, yet comprehensive textbook intended for students interested in gaining a better understanding of the mathematics and statistics that underpin the rich variety of ideas and machine learning algorithms in data science. Key Features: Focuses on mathematical understanding. Presentation is self-contained, accessible, and comprehensive. Extensive list of exercises and worked-out examples. Many concrete algorithms with Python code. Full color throughout. The Authors: Dirk P. Kroese, PhD, is a Professor of Mathematics and Statistics at The University of Queensland. He has published over 120 articles and five books in a wide range of areas in mathematics, statistics, data science, machine learning, and Monte Carlo methods. He is a pioneer of the well-known Cross-Entropy method—an adaptive Monte Carlo technique, which is being used around the world to help solve difficult estimation and optimization problems in science, engineering, and finance. Zdravko Botev, PhD, is an Australian Mathematical Science Institute Lecturer in Data Science and Machine Learning with an appointment at the University of New South Wales in Sydney, Australia. He is the recipient of the 2018 Christopher Heyde Medal of the Australian Academy of Science for distinguished research in the Mathematical Sciences. Thomas Taimre, PhD, is a Senior Lecturer of Mathematics and Statistics at The University of Queensland. His research interests range from applied probability and Monte Carlo methods to applied physics and the remarkably universal self-mixing effect in lasers. He has published over 100 articles, holds a patent, and is the coauthor of Handbook of Monte Carlo Methods (Wiley). Radislav Vaisman, PhD, is a Lecturer of Mathematics and Statistics at The University of Queensland. His research interests lie at the intersection of applied probability, machine learning, and computer science. He has published over 20 articles and two books.

Introducing Data Science

Big data, machine learning, and more, using Python tools *Simon and Schuster*
Summary Introducing Data Science teaches you how to accomplish the fundamental tasks that occupy data scientists. Using the Python language and common Python libraries, you'll experience firsthand the challenges of dealing with data at scale and gain a solid foundation in data science. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Many companies need developers with data science skills to work on projects ranging from social media marketing to machine learning. Discovering what you need to learn to begin a career as a data scientist can seem bewildering. This book is designed to help you get started. About the Book Introducing Data ScienceIntroducing Data Science explains vital data science concepts and teaches you how to accomplish the fundamental tasks that occupy data scientists. You'll explore data visualization, graph databases, the use of NoSQL, and the data science process. You'll use the Python language and common Python libraries as you experience firsthand the challenges of dealing with data at scale. Discover how Python allows you to gain insights from data sets so big that they need to be stored on multiple machines, or from data moving so quickly that no single machine can handle it. This book gives you hands-on experience with the most popular Python data science libraries, Scikit-learn and StatsModels. After reading this book, you'll have the solid foundation you need to start a career in data science. What's Inside Handling large data Introduction to machine learning Using Python to work with data Writing data science algorithms About the Reader This book assumes you're comfortable reading code in Python or a similar language, such as C, Ruby, or JavaScript. No prior experience with data science is required. About the Authors Davy Cielen, Arno D. B. Meysman, and Mohamed Ali are the founders and managing partners of Optimately and Maiton, where they focus on developing data science projects and solutions in various sectors. Table of Contents Data science in a big data world The data science process Machine learning Handling large data on a single computer First steps in big data Join the NoSQL movement The rise of graph databases Text mining and text analytics Data visualization to the end user

Data Scientists at Work *Apress*
Data Scientists at Work is a collection of interviews with sixteen of the world's most influential and innovative data scientists from across the spectrum of this hot new profession. "Data scientist is the sexiest job in the 21st century," according to the Harvard Business Review. By 2018, the United States will experience a shortage of 190,000 skilled data scientists, according to a McKinsey report. Through incisive in-depth interviews, this book mines the what, how, and why of the practice of data science from the stories, ideas, shop talk, and forecasts of its preeminent practitioners across diverse industries: social network (Yann LeCun, Facebook); professional network (Daniel Tunkelang, LinkedIn); venture capital (Roger Ehrenberg, IA Ventures); enterprise cloud computing and neuroscience (Eric Jonas, formerly Salesforce.com); newspaper and media (Chris Wiggins, The New York Times); streaming television (Caitlin Smallwood, Netflix); music forecast (Victor Hu, Next Big Sound); strategic intelligence (Amy

Heineike, Quid); environmental big data (André Karpištšenko, Planet OS); geospatial marketing intelligence (Jonathan Lenaghan, PlaceIQ); advertising (Claudia Perlich, Dstillery); fashion e-commerce (Anna Smith, Rent the Runway); specialty retail (Erin Shellman, Nordstrom); email marketing (John Foreman, MailChimp); predictive sales intelligence (Kira Radinsky, SalesPredict); and humanitarian nonprofit (Jake Porway, DataKind). The book features a stimulating foreword by Google's Director of Research, Peter Norvig. Each of these data scientists shares how he or she tailors the torrent-taming techniques of big data, data visualization, search, and statistics to specific jobs by dint of ingenuity, imagination, patience, and passion. Data Scientists at Work parts the curtain on the interviewees' earliest data projects, how they became data scientists, their discoveries and surprises in working with data, their thoughts on the past, present, and future of the profession, their experiences of team collaboration within their organizations, and the insights they have gained as they get their hands dirty refining mountains of raw data into objects of commercial, scientific, and educational value for their organizations and clients.

Beginning Data Science with Python and Jupyter

Use powerful industry-standard tools within Jupyter and the Python ecosystem to unlock new, actionable insights from your data *Packt Publishing Ltd*

Getting started with data science doesn't have to be an uphill battle. This step-by-step guide is ideal for beginners who know a little Python and are looking for a quick, fast-paced introduction. Key Features Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts like SVM, KNN classifiers and Random Forests Discover how you can use web scraping to gather and parse your own bespoke datasets Book Description Get to grips with the skills you need for entry-level data science in this hands-on Python and Jupyter course. You'll learn about some of the most commonly used libraries that are part of the Anaconda distribution, and then explore machine learning models with real datasets to give you the skills and exposure you need for the real world. We'll finish up by showing you how easy it can be to scrape and gather your own data from the open web, so that you can apply your new skills in an actionable context. What you will learn Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts like SVM, KNN classifiers, and Random Forests Plan a machine learning classification strategy and train classification, models Use validation curves and dimensionality reduction to tune and enhance your models Discover how you can use web scraping to gather and parse your own bespoke datasets Scrape tabular data from web pages and transform them into Pandas DataFrames Create interactive, web-friendly visualizations to clearly communicate your findings Who this book is for This book is ideal for professionals with a variety of job descriptions across large range of industries, given the rising popularity and accessibility of data science. You'll need some prior experience with Python, with any prior work with libraries like Pandas, Matplotlib and Pandas providing you a useful head start.

Approaching (Almost) Any Machine Learning Problem *Abhishek Thakur*

This is not a traditional book. The book has a lot of code. If you don't like the code first approach do not buy this book. Making code available on Github is not an option. This book is for people who have some theoretical knowledge of machine learning and deep learning and want to dive into applied machine learning. The book doesn't explain the algorithms but is more oriented towards how and what should you use to solve machine learning and deep learning problems. The book is not for you if you are looking for pure basics. The book is for you if you are looking for guidance on approaching machine learning problems. The book is best enjoyed with a cup of coffee and a laptop/workstation where you can code along. Table of contents: - Setting up your working environment - Supervised vs unsupervised learning - Cross-validation - Evaluation metrics - Arranging machine learning projects - Approaching categorical variables - Feature engineering - Feature selection - Hyperparameter optimization - Approaching image classification & segmentation - Approaching text classification/regression - Approaching ensembling and stacking - Approaching reproducible code & model serving There are no sub-headings. Important terms are written in bold. I will be answering all your queries related to the book and will be making YouTube tutorials to cover what has not been discussed in the book. To ask questions/doubts, visit this link: <https://bit.ly/aamlquestions> And Subscribe to my youtube channel: <https://bit.ly/abhitubesub>

R Programming for Data Science *Lulu.com*

Data science has taken the world by storm. Every field of study and area of business has been affected as people increasingly realize the value of the incredible quantities of data being generated. But to extract value from those data, one needs to be tra

Building Data Science Applications with FastAPI

Develop, manage, and deploy efficient machine learning applications with Python *Packt Publishing Ltd*

Get well-versed with FastAPI features and best practices for testing, monitoring, and deployment to run high-quality and robust data science applications Key Features Cover the concepts of the FastAPI framework, including aspects relating to asynchronous programming, type hinting, and dependency injection Develop efficient RESTful APIs for data science with modern Python Build, test, and deploy high performing data science and machine learning systems with FastAPI Book Description FastAPI is a web framework for building APIs with Python 3.6 and its later versions based on standard Python-type hints. With this book, you'll be able to create fast and reliable data science API backends using practical examples. This book starts with the basics of the FastAPI framework and associated modern Python programming language concepts. You'll be taken through all the aspects of the framework, including its powerful dependency injection system and how you can use it to communicate with databases, implement authentication and integrate machine learning models. Later, you'll cover best practices relating to testing and deployment to run a high-quality and robust application. You'll also be introduced to the extensive ecosystem of Python data science packages. As you progress, you'll learn how to build data science applications in Python using FastAPI. The book also demonstrates how to develop fast and efficient machine learning prediction

backends and test them to achieve the best performance. Finally, you'll see how to implement a real-time face detection system using WebSockets and a web browser as a client. By the end of this FastAPI book, you'll have not only learned how to implement Python in data science projects but also how to maintain and design them to meet high programming standards with the help of FastAPI. What you will learn Explore the basics of modern Python and async I/O programming Get to grips with basic and advanced concepts of the FastAPI framework Implement a FastAPI dependency to efficiently run a machine learning model Integrate a simple face detection algorithm in a FastAPI backend Integrate common Python data science libraries in a web backend Deploy a performant and reliable web backend for a data science application Who this book is for This Python data science book is for data scientists and software developers interested in gaining knowledge of FastAPI and its ecosystem to build data science applications. Basic knowledge of data science and machine learning concepts and how to apply them in Python is recommended.

Data Science and Analytics with Python *CRC Press*

Data Science and Analytics with Python is designed for practitioners in data science and data analytics in both academic and business environments. The aim is to present the reader with the main concepts used in data science using tools developed in Python, such as SciKit-learn, Pandas, Numpy, and others. The use of Python is of particular interest, given its recent popularity in the data science community. The book can be used by seasoned programmers and newcomers alike. The book is organized in a way that individual chapters are sufficiently independent from each other so that the reader is comfortable using the contents as a reference. The book discusses what data science and analytics are, from the point of view of the process and results obtained. Important features of Python are also covered, including a Python primer. The basic elements of machine learning, pattern recognition, and artificial intelligence that underpin the algorithms and implementations used in the rest of the book also appear in the first part of the book. Regression analysis using Python, clustering techniques, and classification algorithms are covered in the second part of the book. Hierarchical clustering, decision trees, and ensemble techniques are also explored, along with dimensionality reduction techniques and recommendation systems. The support vector machine algorithm and the Kernel trick are discussed in the last part of the book. About the Author Dr. Jesús Rogel-Salazar is a Lead Data scientist with experience in the field working for companies such as AKQA, IBM Data Science Studio, Dow Jones and others. He is a visiting researcher at the Department of Physics at Imperial College London, UK and a member of the School of Physics, Astronomy and Mathematics at the University of Hertfordshire, UK, He obtained his doctorate in physics at Imperial College London for work on quantum atom optics and ultra-cold matter. He has held a position as senior lecturer in mathematics as well as a consultant in the financial industry since 2006. He is the author of the book Essential Matlab and Octave, also published by CRC Press. His interests include mathematical modelling, data science, and optimization in a wide range of applications including optics, quantum mechanics, data journalism, and finance.

Data Science for Business

What You Need to Know about Data Mining and Data-Analytic Thinking *"O'Reilly Media, Inc."*

Written by renowned data science experts Foster Provost and Tom Fawcett, Data Science for Business introduces the fundamental principles of data science, and walks you through the "data-analytic thinking" necessary for extracting useful knowledge and business value from the data you collect. This guide also helps you understand the many data-mining techniques in use today. Based on an MBA course Provost has taught at New York University over the past ten years, Data Science for Business provides examples of real-world business problems to illustrate these principles. You'll not only learn how to improve communication between business stakeholders and data scientists, but also how participate intelligently in your company's data science projects. You'll also discover how to think data-analytically, and fully appreciate how data science methods can support business decision-making. Understand how data science fits in your organization—and how you can use it for competitive advantage Treat data as a business asset that requires careful investment if you're to gain real value Approach business problems data-analytically, using the data-mining process to gather good data in the most appropriate way Learn general concepts for actually extracting knowledge from data Apply data science principles when interviewing data science job candidates

Data Science for Business With R *SAGE Publications, Incorporated*

Data Science for Business with R, written by Jeffrey S. Saltz and Jeffrey M. Stanton, focuses on the concepts foundational for students starting a business analytics or data science degree program. To keep the book practical and applied, the authors feature a running case using a global airline business's customer survey dataset to illustrate how to turn data in business decisions, in addition to numerous examples throughout. To aid in usability beyond the classroom, the text features full integration of freely-available R and RStudio software, one of the most popular data science tools available. Designed for students with little to no experience in related areas like computer science, the book chapters follow a logical order from introduction and installation of R and RStudio, working with data architecture, undertaking data collection, performing data analysis, and transitioning to data archiving and presentation. Each chapter follows a familiar structure, starting with learning objectives and background, following the basic steps of functions alongside simple examples, applying these functions to the case study, and ending with chapter challenge questions, sources, and a list of R functions so students know what to expect in each step of their data science course. Data Science for Business with R provides readers with a straightforward and applied guide to this new and evolving field.

Basketball Data Science

With Applications in R *CRC Press*

Using data from one season of NBA games, Basketball Data Science: With Applications in R is the perfect book for anyone interested in learning and applying data analytics in basketball. Whether assessing the spatial performance of an MBA player's shots or doing an analysis of the impact of high

pressure game situations on the probability of scoring, this book discusses a variety of case studies and hands-on examples using a custom R package. The codes are supplied so readers can reproduce the analyses themselves or create their own. Assuming a basic statistical knowledge, Basketball Data Science with R is suitable for students, technicians, coaches, data analysts and applied researchers. Features: · One of the first books to provide statistical and data mining methods for the growing field of analytics in basketball. · Presents tools for modelling graphs and figures to visualize the data. · Includes real world case studies and examples, such as estimations of scoring probability using the Golden State Warriors as a test case. · Provides the source code and data so readers can do their own analyses on NBA teams and players.

Doing Data Science

Straight Talk from the Frontline "O'Reilly Media, Inc."

Now that people are aware that data can make the difference in an election or a business model, data science as an occupation is gaining ground. But how can you get started working in a wide-ranging, interdisciplinary field that’s so clouded in hype? This insightful book, based on Columbia University’s Introduction to Data Science class, tells you what you need to know. In many of these chapter-long lectures, data scientists from companies such as Google, Microsoft, and eBay share new algorithms, methods, and models by presenting case studies and the code they use. If you’re familiar with linear algebra, probability, and statistics, and have programming experience, this book is an ideal introduction to data science. Topics include: Statistical inference, exploratory data analysis, and the data science process Algorithms Spam filters, Naive Bayes, and data wrangling Logistic regression Financial modeling Recommendation engines and causality Data visualization Social networks and data journalism Data engineering, MapReduce, Pregel, and Hadoop Doing Data Science is collaboration between course instructor Rachel Schutt, Senior VP of Data Science at News Corp, and data science consultant Cathy O’Neil, a senior data scientist at Johnson Research Labs, who attended and blogged about the course.

Getting Started with R Oxford University Press

R is rapidly becoming the standard software for statistical analyses, graphical presentation of data, and programming in the natural, physical, social, and engineering sciences. Getting Started with R is now the go-to introductory guide for biologists wanting to learn how to use R in their research. It teaches readers how to import, explore, graph, and analyse data, while keeping them focused on their ultimate goals: clearly communicating their data in oral presentations, posters, papers, and reports. It provides a consistent workflow for using R that is simple, efficient, reliable, and reproducible. This second edition has been updated and expanded while retaining the concise and engaging nature of its predecessor, offering an accessible and fun introduction to the packages dplyr and ggplot2 for data manipulation and graphing. It expands the set of basic statistics considered in the first edition to include new examples of a simple regression, a one-way and a two-way ANOVA. Finally, it introduces a new chapter on the generalised linear model. Getting Started with R is suitable for undergraduates, graduate students, professional researchers, and practitioners in the biological sciences.

Data Science Bookcamp

Five real-world Python projects Simon and Schuster

Learn data science with Python by building five real-world projects! Experiment with card game predictions, tracking disease outbreaks, and more, as you build a flexible and intuitive understanding of data science. In Data Science Bookcamp you will learn: - Techniques for computing and plotting probabilities - Statistical analysis using Scipy - How to organize datasets with clustering algorithms - How to visualize complex multi-variable datasets

- How to train a decision tree machine learning algorithm In Data Science Bookcamp you’ll test and build your knowledge of Python with the kind of open-ended problems that professional data scientists work on every day. Downloadable data sets and thoroughly-explained solutions help you lock in what you’ve learned, building your confidence and making you ready for an exciting new data science career. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology A data science project has a lot of moving parts, and it takes practice and skill to get all the code, algorithms, datasets, formats, and visualizations working together harmoniously. This unique book guides you through five realistic projects, including tracking disease outbreaks from news headlines, analyzing social networks, and finding relevant patterns in ad click data. About the book Data Science Bookcamp doesn’t stop with surface-level theory and toy examples. As you work through each project, you’ll learn how to troubleshoot common problems like missing data, messy data, and algorithms that don’t quite fit the model you’re building. You’ll appreciate the detailed setup instructions and the fully explained solutions that highlight common failure points. In the end, you’ll be confident in your skills because you can see the results. What's inside - Web scraping - Organize datasets with clustering algorithms - Visualize complex multi-variable datasets - Train a decision tree machine learning algorithm About the reader For readers who know the basics of Python. No prior data science or machine learning skills required. About the author Leonard Apeltsin is the Head of Data Science at Anomaly, where his team applies advanced analytics to uncover healthcare fraud, waste, and abuse. Table of Contents CASE STUDY 1 FINDING THE WINNING STRATEGY IN A CARD GAME 1 Computing probabilities using Python 2 Plotting probabilities using Matplotlib 3 Running random simulations in NumPy 4 Case study 1 solution CASE STUDY 2 ASSESSING ONLINE AD CLICKS FOR SIGNIFICANCE 5 Basic probability and statistical analysis using SciPy 6 Making predictions using the central limit theorem and SciPy 7 Statistical hypothesis testing 8 Analyzing tables using Pandas 9 Case study 2 solution CASE STUDY 3 TRACKING DISEASE OUTBREAKS USING NEWS HEADLINES 10 Clustering data into groups 11 Geographic location visualization and analysis 12 Case study 3 solution CASE STUDY 4 USING ONLINE JOB POSTINGS TO IMPROVE YOUR DATA SCIENCE RESUME 13 Measuring text similarities 14 Dimension reduction of matrix data 15 NLP analysis of large text datasets 16 Extracting text from web pages 17 Case study 4 solution CASE STUDY 5 PREDICTING FUTURE FRIENDSHIPS FROM SOCIAL NETWORK DATA 18 An introduction to graph theory and network analysis 19 Dynamic graph theory techniques for node ranking and social network analysis 20 Network-driven supervised machine learning 21 Training linear classifiers with logistic regression 22 Training nonlinear classifiers with decision tree techniques 23 Case study 5 solution

Becoming a Data Head

How to Think, Speak, and Understand Data Science, Statistics, and Machine Learning John Wiley & Sons

"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.