

Python Data Science Essentials Learn The Fundamentals Of Data Science With Python

Getting the books **Python Data Science Essentials Learn The Fundamentals Of Data Science With Python** now is not type of challenging means. You could not lonesome going in the same way as ebook addition or library or borrowing from your associates to entre them. This is an definitely easy means to specifically acquire guide by on-line. This online proclamation Python Data Science Essentials Learn The Fundamentals Of Data Science With Python can be one of the options to accompany you as soon as having extra time.

It will not waste your time. say yes me, the e-book will unconditionally proclaim you additional business to read. Just invest little period to entrance this on-line pronouncement **Python Data Science Essentials Learn The Fundamentals Of Data Science With Python** as with ease as review them wherever you are now.

[Artificial Intelligence with Python](#) - Prateek Joshi 2017-01-27

Build real-world Artificial Intelligence applications with Python to intelligently interact with the world around you About This Book Step into the amazing world of intelligent apps using this comprehensive guide Enter the world of Artificial Intelligence, explore it, and create your own applications Work through simple yet insightful examples that will get you up and running with Artificial Intelligence in no time Who This Book Is For This book is for Python developers who want to build real-world Artificial Intelligence applications. This book is friendly to Python beginners, but being familiar with Python would be useful to play around with the code. It will also be useful for experienced Python programmers who are looking to use Artificial Intelligence techniques in their existing technology stacks. What You Will Learn Realize different classification and regression techniques Understand the concept of clustering and how to use it to automatically segment data See how to build an intelligent recommender system Understand logic programming and how to use it Build automatic speech recognition systems Understand the basics of heuristic search and genetic programming Develop games using Artificial Intelligence Learn how reinforcement learning works Discover how to build intelligent applications centered on images, text, and time series data See how to use deep learning algorithms and build applications based on it In Detail Artificial Intelligence is becoming increasingly relevant in the modern world where everything is driven by technology and data. It is used extensively across many fields such as search engines, image recognition, robotics, finance, and so on. We will explore various real-world scenarios in this book and you'll learn about various algorithms that can be used to build Artificial Intelligence applications. During the course of this book, you will find out how to make informed decisions about what algorithms to use in a given context. Starting from the basics of Artificial Intelligence, you will learn how to develop various building blocks using different data mining techniques. You will see how to implement different algorithms to get the best possible results, and will understand how to apply them to real-world scenarios. If you want to add an intelligence layer to any application that's based on images, text, stock market, or some other form of data, this exciting book on Artificial Intelligence will definitely be your guide! Style and approach This highly practical book will show you how to implement Artificial Intelligence. The book provides multiple examples enabling you to create smart applications to meet the needs of your organization. In every chapter, we explain an algorithm, implement it, and then build a smart application.

Machine Learning - Samuel Hack 2021-01-07

Master the world of Python and Machine Learning with this incredible 4-in-1 bundle. Are you interested in becoming a Python pro? Do you want to learn more about the incredible world of machine learning, and what it can do for you? Then keep reading. Created with the beginner in mind, this powerful bundle delves into the fundamentals behind Python and Machine Learning, from basic code and mathematical formulas to complex neural networks and ensemble modeling. Inside, you'll discover everything you need to know to get started with Python and Machine Learning, and begin your journey to success! In book one - MACHINE LEARNING FOR BEGINNERS, you'll learn: What is Artificial Intelligence Really, and Why is it So Powerful? Choosing the Right Kind of Machine Learning Model for You An Introduction to Statistics Reinforcement Learning and Ensemble Modeling "Random Forests" and Decision Trees In book two - MACHINE LEARNING MATHEMATICS, you will: Learn the Fundamental Concepts of Machine Learning Algorithms Understand The Four Fundamental Types of Machine Learning Algorithm Master the Concept of "Statistical Learning" Learn Everything You Need to Know

about Neural Networks and Data Pipelines Master the Concept of "General Setting of Learning" In book three - LEARNING PYTHON, you'll discover: How to Install, Run, and Understand Python on Any Operating System A Comprehensive Introduction to Python Python Basics and Writing Code Writing Loops, Conditional Statements, Exceptions and More Python Expressions and The Beauty of Inheritances And in book four - PYTHON MACHINE LEARNING, you will: Learn the Fundamentals of Machine Learning Master the Nuances of 12 of the Most Popular and Widely-Used Machine Learning Algorithms Become Familiar with Data Science Technology Dive Into the Functioning of Scikit-Learn Library and Develop Machine Learning Models Uncover the Secrets of the Most Critical Aspect of Developing a Machine Learning Model - Data Pre-Processing and Training/Testing Subsets Whether you're a complete beginner or a programmer looking to improve your skillset, this bundle is your all-in-one solution to mastering the world of Python and Machine Learning. So don't wait - it's never been easier to learn. Buy Now to Become a Master of Python and Machine Learning Today!

[Python Machine Learning](#) - Samuel Hack 2019-08-28

Buy the Paperback Version of this Book and get the Kindle Book Version for FREE Discover the incredible world of Machine Learning with this amazing guide Do you want to understand machine learning, but it all looks too daunting and complex? Afraid to open the 'pandora's box' and waste hours searching for answers? Then keep reading Written with the beginner in mind, this powerful guide breaks down everything you need to know about machine learning and Python in a simple, easy-to-understand way. So many other books make machine learning look impossible to understand and even harder to master - but now you can familiarize yourself with this incredible technology like never before! With a detailed and concise overview of the fundamentals, along with the challenges and limitations currently being tackled by the pros, inside this comprehensive guide you will: Learn the Fundamentals of Machine Learning which Are Being Developed and Advanced with Python Master the Nuances of 12 of the Most Popular and Widely-Used Machine Learning Algorithms, in a Language that Requires No Prior Background in Python Discover the Details of the Supervised, Unsupervised, and Reinforcement Algorithms, which Serve as the Skeleton of Hundreds of Machine Learning Algorithms Being Developed Every Day Become Familiar with Data Science Technology, an Umbrella Term Used for the Cutting-Edge Technologies of Today Dive Into the Functioning of Scikit-Learn Library and Develop Machine Learning Models, with a Detailed Walkthrough and Open Source Database using Illustrations and actual Python Code Understand the Entire Process of Creating Neural Network Models on TensorFlow, Using Open Source Data Sets and real Python Code Uncover the Secrets of the Most Critical Aspect of Developing a Machine Learning Model - Data Pre-Processing and Training/Testing Subsets And So Much More! With a wealth of tips and tricks, along with invaluable advice guaranteed to help you with your machine learning journey, this book is a powerful and revolutionary tool for creating, developing, and using machine learning. From understanding the Python language to creating data sets and building neural networks, now you can become the master of machine learning with this incredible guide! So what are you waiting for? Buy now and join the millions of people using machine learning today! Scroll Up and Click the "Buy now" Button!

[An Introduction to Statistical Learning](#) - Gareth James 2013-06-24

An Introduction to Statistical Learning provides an accessible overview of the field of statistical learning, an essential toolset for making sense of the vast and complex data sets that have emerged in fields ranging from biology to finance to marketing to astrophysics in the past twenty years. This book presents some of the most important modeling and prediction

techniques, along with relevant applications. Topics include linear regression, classification, resampling methods, shrinkage approaches, tree-based methods, support vector machines, clustering, and more. Color graphics and real-world examples are used to illustrate the methods presented. Since the goal of this textbook is to facilitate the use of these statistical learning techniques by practitioners in science, industry, and other fields, each chapter contains a tutorial on implementing the analyses and methods presented in R, an extremely popular open source statistical software platform. Two of the authors co-wrote *The Elements of Statistical Learning* (Hastie, Tibshirani and Friedman, 2nd edition 2009), a popular reference book for statistics and machine learning researchers. An *Introduction to Statistical Learning* covers many of the same topics, but at a level accessible to a much broader audience. This book is targeted at statisticians and non-statisticians alike who wish to use cutting-edge statistical learning techniques to analyze their data. The text assumes only a previous course in linear regression and no knowledge of matrix algebra.

Mastering Python for Data Science - Samir Madhavan 2015-08-31

Explore the world of data science through Python and learn how to make sense of data About This Book Master data science methods using Python and its libraries Create data visualizations and mine for patterns Advanced techniques for the four fundamentals of Data Science with Python - data mining, data analysis, data visualization, and machine learning Who This Book Is For If you are a Python developer who wants to master the world of data science then this book is for you. Some knowledge of data science is assumed. What You Will Learn Manage data and perform linear algebra in Python Derive inferences from the analysis by performing inferential statistics Solve data science problems in Python Create high-end visualizations using Python Evaluate and apply the linear regression technique to estimate the relationships among variables. Build recommendation engines with the various collaborative filtering algorithms Apply the ensemble methods to improve your predictions Work with big data technologies to handle data at scale In Detail Data science is a relatively new knowledge domain which is used by various organizations to make data driven decisions. Data scientists have to wear various hats to work with data and to derive value from it. The Python programming language, beyond having conquered the scientific community in the last decade, is now an indispensable tool for the data science practitioner and a must-know tool for every aspiring data scientist. Using Python will offer you a fast, reliable, cross-platform, and mature environment for data analysis, machine learning, and algorithmic problem solving. This comprehensive guide helps you move beyond the hype and transcend the theory by providing you with a hands-on, advanced study of data science. Beginning with the essentials of Python in data science, you will learn to manage data and perform linear algebra in Python. You will move on to deriving inferences from the analysis by performing inferential statistics, and mining data to reveal hidden patterns and trends. You will use the matplotlib library to create high-end visualizations in Python and uncover the fundamentals of machine learning. Next, you will apply the linear regression technique and also learn to apply the logistic regression technique to your applications, before creating recommendation engines with various collaborative filtering algorithms and improving your predictions by applying the ensemble methods. Finally, you will perform K-means clustering, along with an analysis of unstructured data with different text mining techniques and leveraging the power of Python in big data analytics. Style and approach This book is an easy-to-follow, comprehensive guide on data science using Python. The topics covered in the book can all be used in real world scenarios.

R for Data Science - Hadley Wickham 2016-12-12

Learn how to use R to turn raw data into insight, knowledge, and understanding. This book introduces you to R, RStudio, and the tidyverse, a collection of R packages designed to work together to make data science fast, fluent, and fun. Suitable for readers with no previous programming experience, *R for Data Science* is designed to get you doing data science as quickly as possible. Authors Hadley Wickham and Garrett Golemund guide you through the steps of importing, wrangling, exploring, and modeling your data and communicating the results. You'll get a complete, big-picture understanding of the data science cycle, along with basic tools you need to manage the details. Each section of the book is paired with exercises to help you practice what you've learned along the way. You'll learn how to: Wrangle—transform your datasets into a form convenient for analysis Program—learn powerful R tools for solving data problems with greater clarity and ease Explore—examine your data, generate hypotheses, and quickly test them

Model—provide a low-dimensional summary that captures true "signals" in your dataset Communicate—learn R Markdown for integrating prose, code, and results

Python Data Science Handbook - Jake VanderPlas 2016-11-21

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

Learn Ethical Hacking from Scratch - Zaid Sabih 2018-07-31

Learn how to hack systems like black hat hackers and secure them like security experts Key Features Understand how computer systems work and their vulnerabilities Exploit weaknesses and hack into machines to test their security Learn how to secure systems from hackers Book Description This book starts with the basics of ethical hacking, how to practice hacking safely and legally, and how to install and interact with Kali Linux and the Linux terminal. You will explore network hacking, where you will see how to test the security of wired and wireless networks. You'll also learn how to crack the password for any Wi-Fi network (whether it uses WEP, WPA, or WPA2) and spy on the connected devices. Moving on, you will discover how to gain access to remote computer systems using client-side and server-side attacks. You will also get the hang of post-exploitation techniques, including remotely controlling and interacting with the systems that you compromised. Towards the end of the book, you will be able to pick up web application hacking techniques. You'll see how to discover, exploit, and prevent a number of website vulnerabilities, such as XSS and SQL injections. The attacks covered are practical techniques that work against real systems and are purely for educational purposes. At the end of each section, you will learn how to detect, prevent, and secure systems from these attacks. What you will learn Understand ethical hacking and the different fields and types of hackers Set up a penetration testing lab to practice safe and legal hacking Explore Linux basics, commands, and how to interact with the terminal Access password-protected networks and spy on connected clients Use server and client-side attacks to hack and control remote computers Control a hacked system remotely and use it to hack other systems Discover, exploit, and prevent a number of web application vulnerabilities such as XSS and SQL injections Who this book is for Learning *Ethical Hacking from Scratch* is for anyone interested in learning how to hack and test the security of systems like professional hackers and security experts.

Build a Career in Data Science - Emily Robinson 2020-03-06

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

Practical Data Science with Python - Nathan George 2021-09-30

Learn to effectively manage data and execute data science projects from start to finish using Python Key Features Understand and utilize data science tools in Python, such as specialized machine learning algorithms and statistical modeling Build a strong data science foundation with the best data science tools available in Python Add value to yourself, your organization, and society by extracting actionable insights from raw data Book Description Practical Data Science with Python teaches you core data science concepts, with real-world and realistic examples, and strengthens your grip on the basic as well as advanced principles of data preparation and storage, statistics, probability theory, machine learning, and Python programming, helping you build a solid foundation to gain proficiency in data science. The book starts with an overview of basic Python skills and then introduces foundational data science techniques, followed by a thorough explanation of the Python code needed to execute the techniques. You'll understand the code by working through the examples. The code has been broken down into small chunks (a few lines or a function at a time) to enable thorough discussion. As you progress, you will learn how to perform data analysis while exploring the functionalities of key data science Python packages, including pandas, SciPy, and scikit-learn. Finally, the book covers ethics and privacy concerns in data science and suggests resources for improving data science skills, as well as ways to stay up to date on new data science developments. By the end of the book, you should be able to comfortably use Python for basic data science projects and should have the skills to execute the data science process on any data source. What you will learn Use Python data science packages effectively Clean and prepare data for data science work, including feature engineering and feature selection Data modeling, including classic statistical models (such as t-tests), and essential machine learning algorithms, such as random forests and boosted models Evaluate model performance Compare and understand different machine learning methods Interact with Excel spreadsheets through Python Create automated data science reports through Python Get to grips with text analytics techniques Who this book is for The book is intended for beginners, including students starting or about to start a data science, analytics, or related program (e.g. Bachelor's, Master's, bootcamp, online courses), recent college graduates who want to learn new skills to set them apart in the job market, professionals who want to learn hands-on data science techniques in Python, and those who want to shift their career to data science. The book requires basic familiarity with Python. A "getting started with Python" section has been included to get complete novices up to speed.

Data Science - John D. Kelleher 2018-04-13

A concise introduction to the emerging field of data science, explaining its evolution, relation to machine learning, current uses, data infrastructure issues, and ethical challenges. The goal of data science is to improve decision making through the analysis of data. Today data science determines the ads we see online, the books and movies that are recommended to us online, which emails are filtered into our spam folders, and even how much we pay for health insurance. This volume in the MIT Press Essential Knowledge series offers a concise introduction to the emerging field of data science, explaining its evolution, current uses, data infrastructure issues, and ethical challenges. It has never been easier for organizations to gather, store, and process data. Use of data science is driven by the rise of big data and social media, the development of high-performance computing, and the emergence of such powerful methods for data analysis and modeling as deep learning. Data science encompasses a set of principles, problem definitions, algorithms, and processes for extracting non-obvious and useful patterns from large datasets. It is closely related to the fields of data mining and machine

learning, but broader in scope. This book offers a brief history of the field, introduces fundamental data concepts, and describes the stages in a data science project. It considers data infrastructure and the challenges posed by integrating data from multiple sources, introduces the basics of machine learning, and discusses how to link machine learning expertise with real-world problems. The book also reviews ethical and legal issues, developments in data regulation, and computational approaches to preserving privacy. Finally, it considers the future impact of data science and offers principles for success in data science projects.

Fundamentals of Data Visualization - Claus O. Wilke 2019-03-18

Effective visualization is the best way to communicate information from the increasingly large and complex datasets in the natural and social sciences. But with the increasing power of visualization software today, scientists, engineers, and business analysts often have to navigate a bewildering array of visualization choices and options. This practical book takes you through many commonly encountered visualization problems, and it provides guidelines on how to turn large datasets into clear and compelling figures. What visualization type is best for the story you want to tell? How do you make informative figures that are visually pleasing? Author Claus O. Wilke teaches you the elements most critical to successful data visualization. Explore the basic concepts of color as a tool to highlight, distinguish, or represent a value Understand the importance of redundant coding to ensure you provide key information in multiple ways Use the book's visualizations directory, a graphical guide to commonly used types of data visualizations Get extensive examples of good and bad figures Learn how to use figures in a document or report and how employ them effectively to tell a compelling story

Introduction to Data Science - Laura Igual 2017-02-22

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

Data Science and Machine Learning - Dirk P. Kroese 2019-11-20

"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. Further Resources can be found on the authors website: <https://github.com/DSML-book/Lectures>

Python: Real-World Data Science - Dusty Phillips 2016-06-10

Unleash the power of Python and its robust data science capabilities About This Book Unleash the power of Python 3 objects Learn to use powerful Python libraries for effective data processing and analysis Harness the power of Python to analyze data and create insightful predictive models Unlock deeper insights into machine learning with this vital guide to cutting-edge predictive analytics Who This Book Is For

Entry-level analysts who want to enter in the data science world will find this course very useful to get themselves acquainted with Python's data science capabilities for doing real-world data analysis. What You Will Learn Install and setup Python Implement objects in Python by creating classes and defining methods Get acquainted with NumPy to use it with arrays and array-oriented computing in data analysis Create effective visualizations for presenting your data using Matplotlib Process and analyze data using the time series capabilities of pandas Interact with different kind of database systems, such as file, disk format, Mongo, and Redis Apply data mining concepts to real-world problems Compute on big data, including real-time data from the Internet Explore how to use different machine learning models to ask different questions of your data In Detail The Python: Real-World Data Science course will take you on a journey to become an efficient data science practitioner by thoroughly understanding the key concepts of Python. This learning path is divided into four modules and each module are a mini course in their own right, and as you complete each one, you'll have gained key skills and be ready for the material in the next module. The course begins with getting your Python fundamentals nailed down. After getting familiar with Python core concepts, it's time that you dive into the field of data science. In the second module, you'll learn how to perform data analysis using Python in a practical and example-driven way. The third module will teach you how to design and develop data mining applications using a variety of datasets, starting with basic classification and affinity analysis to more complex data types including text, images, and graphs. Machine learning and predictive analytics have become the most important approaches to uncover data gold mines. In the final module, we'll discuss the necessary details regarding machine learning concepts, offering intuitive yet informative explanations on how machine learning algorithms work, how to use them, and most importantly, how to avoid the common pitfalls. Style and approach This course includes all the resources that will help you jump into the data science field with Python and learn how to make sense of data. The aim is to create a smooth learning path that will teach you how to get started with powerful Python libraries and perform various data science techniques in depth.

Data Science for Business Professionals - Probyto Data Science and Consulting Pvt. Ltd. 2020-05-06

Primer into the multidisciplinary world of Data Science KEY FEATURES - Explore and use the key concepts of Statistics required to solve data science problems - Use Docker, Jenkins, and Git for Continuous Development and Continuous Integration of your web app - Learn how to build Data Science solutions with GCP and AWS DESCRIPTION The book will initially explain the What-Why of Data Science and the process of solving a Data Science problem. The fundamental concepts of Data Science, such as Statistics, Machine Learning, Business Intelligence, Data pipeline, and Cloud Computing, will also be discussed. All the topics will be explained with an example problem and will show how the industry approaches to solve such a problem. The book will pose questions to the learners to solve the problems and build the problem-solving aptitude and effectively learn. The book uses Mathematics wherever necessary and will show you how it is implemented using Python with the help of an example dataset. WHAT WILL YOU LEARN - Understand the multi-disciplinary nature of Data Science - Get familiar with the key concepts in Mathematics and Statistics - Explore a few key ML algorithms and their use cases - Learn how to implement the basics of Data Pipelines - Get an overview of Cloud Computing & DevOps - Learn how to create visualizations using Tableau WHO THIS BOOK IS FOR This book is ideal for Data Science enthusiasts who want to explore various aspects of Data Science. Useful for Academicians, Business owners, and Researchers for a quick reference on industrial practices in Data Science. TABLE OF CONTENTS 1. Data Science in Practice 2. Mathematics Essentials 3. Statistics Essentials 4. Exploratory Data Analysis 5. Data preprocessing 6. Feature Engineering 7. Machine learning algorithms 8. Productionizing ML models 9. Data Flows in Enterprises 10. Introduction to Databases 11. Introduction to Big Data 12. DevOps for Data Science 13. Introduction to Cloud Computing 14. Deploy Model to Cloud 15. Introduction to Business Intelligence 16. Data Visualization Tools 17. Industry Use Case 1 - FormAssist 18. Industry Use Case 2 - PeopleReporter 19. Data Science Learning Resources 20. Do It Your Self Challenges 21. MCQs for Assessments

Data Science for Business - Foster Provost 2013-07-27

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

Mastering Python for Data Science - Samir Medhaven 2015-08-31

Explore the world of data science through Python and learn how to make sense of data About This Book • Master data science methods using Python and its libraries • Create data visualizations and mine for patterns • Advanced techniques for the four fundamentals of Data Science with Python - data mining, data analysis, data visualization, and machine learning Who This Book Is For If you are a Python developer who wants to master the world of data science then this book is for you. Some knowledge of data science is assumed. What You Will Learn • Manage data and perform linear algebra in Python • Derive inferences from the analysis by performing inferential statistics • Solve data science problems in Python • Create high-end visualizations using Python • Evaluate and apply the linear regression technique to estimate the relationships among variables. • Build recommendation engines with the various collaborative filtering algorithms • Apply the ensemble methods to improve your predictions • Work with big data technologies to handle data at scale In Detail Data science is a relatively new knowledge domain which is used by various organizations to make data driven decisions. Data scientists have to wear various hats to work with data and to derive value from it. The Python programming language, beyond having conquered the scientific community in the last decade, is now an indispensable tool for the data science practitioner and a must-know tool for every aspiring data scientist. Using Python will offer you a fast, reliable, cross-platform, and mature environment for data analysis, machine learning, and algorithmic problem solving. This comprehensive guide helps you move beyond the hype and transcend the theory by providing you with a hands-on, advanced study of data science. Beginning with the essentials of Python in data science, you will learn to manage data and perform linear algebra in Python. You will move on to deriving inferences from the analysis by performing inferential statistics, and mining data to reveal hidden patterns and trends. You will use the matplotlib library to create high-end visualizations in Python and uncover the fundamentals of machine learning. Next, you will apply the linear regression technique and also learn to apply the logistic regression technique to your applications, before creating recommendation engines with various collaborative filtering algorithms and improving your predictions by applying the ensemble methods. Finally, you will perform K-means clustering, along with an analysis of unstructured data with different text mining techniques and leveraging the power of Python in big data analytics. Style and approach This book is an easy-to-follow, comprehensive guide on data science using Python. The topics covered in the book can all be used in real world scenarios.

Mining Social Media - Lam Thuy Vo 2019-11-25

BuzzFeed News Senior Reporter Lam Thuy Vo explains how to mine, process, and analyze data from the social web in meaningful ways with the Python programming language. Did fake Twitter accounts help sway a presidential election? What can Facebook and Reddit archives tell us about human behavior? In Mining Social Media, senior BuzzFeed reporter Lam Thuy Vo shows you how to use Python and key data analysis tools to find the stories buried in social media. Whether you're a professional journalist, an academic researcher, or a citizen investigator, you'll learn how to use technical tools to collect and analyze data from social media sources to build compelling, data-driven stories. Learn how to: Write Python scripts and use APIs to gather data from the social web Download data archives and dig through them for insights Inspect HTML downloaded from websites for useful content Format, aggregate, sort, and filter your collected data using Google Sheets Create data visualizations to illustrate your discoveries Perform advanced data analysis using Python, Jupyter Notebooks, and the pandas library Apply

what you've learned to research topics on your own Social media is filled with thousands of hidden stories just waiting to be told. Learn to use the data-sleuthing tools that professionals use to write your own data-driven stories.

Statistics for Beginners in Data Science - Ai Publishing 2020-04-18
Statistics for Beginners in Data Science Statistical methods are an integral part of data science. Hence, a formal training in statistics is indispensable for data scientists. If you are keen on getting your foot into the lucrative data science and analysis universe, you need to have a fundamental understanding of statistical analysis. Besides, Python is a versatile programming language you need to master to become a career data scientist. As a data scientist, you will identify, clean, explore, analyze, and interpret trends or possible patterns in complex data sets. The explosive growth of Big Data means you have to manage enormous amounts of data, clean it, manipulate it, and process it. Only then the most relevant data can be used. Python is a natural data science tool as it has an assortment of useful libraries, such as Pandas, NumPy, SciPy, Matplotlib, Seaborn, StatsModels, IPython, and several more. And Python's focus on simplicity makes it relatively easy for you to learn. Importantly, the ease of performing repetitive tasks saves you precious time. Long story short--Python is simply a high-priority data science tool. How Is This Book Different? The book focuses equally on the theoretical as well as practical aspects of data science. You will learn how to implement elementary data science tools and algorithms from scratch. The book contains an in-depth theoretical and analytical explanation of all data science concepts and also includes dozens of hands-on, real-life projects that will help you understand the concepts better. The ready-to-access Python codes at various places right through the book are aimed at shortening your learning curve. The main goal is to present you with the concepts, the insights, the inspiration, and the right tools needed to dive into coding and analyzing data in Python. The main benefit of purchasing this book is you get quick access to all the extra content provided with this book--Python codes, exercises, references, and PDFs--on the publisher's website, at no extra price. You get to experiment with the practical aspects of Data Science right from page 1. Beginners in Python and statistics will find this book extremely informative, practical, and helpful. Even if you aren't new to Python and data science, you'll find the hands-on projects in this book immensely helpful. The topics covered include: Introduction to Statistics Getting Familiar with Python Data Exploration and Data Analysis Pandas, Matplotlib, and Seaborn for Statistical Visualization Exploring Two or More Variables and Categorical Data Statistical Tests and ANOVA Confidence Interval Regression Analysis Classification Analysis Click the BUY button and download the book now to start learning and coding Python for Data Science.

Hands-on Scikit-Learn for Machine Learning Applications - David Paper 2019-11-16

Aspiring data science professionals can learn the Scikit-Learn library along with the fundamentals of machine learning with this book. The book combines the Anaconda Python distribution with the popular Scikit-Learn library to demonstrate a wide range of supervised and unsupervised machine learning algorithms. Care is taken to walk you through the principles of machine learning through clear examples written in Python that you can try out and experiment with at home on your own machine. All applied math and programming skills required to master the content are covered in this book. In-depth knowledge of object-oriented programming is not required as working and complete examples are provided and explained. Coding examples are in-depth and complex when necessary. They are also concise, accurate, and complete, and complement the machine learning concepts introduced. Working the examples helps to build the skills necessary to understand and apply complex machine learning algorithms. Hands-on Scikit-Learn for Machine Learning Applications is an excellent starting point for those pursuing a career in machine learning. Students of this book will learn the fundamentals that are a prerequisite to competency. Readers will be exposed to the Anaconda distribution of Python that is designed specifically for data science professionals, and will build skills in the popular Scikit-Learn library that underlies many machine learning applications in the world of Python. What You'll Learn Work with simple and complex datasets common to Scikit-Learn Manipulate data into vectors and matrices for algorithmic processing Become familiar with the Anaconda distribution used in data science Apply machine learning with Classifiers, Regressors, and Dimensionality Reduction Tune algorithms and find the best algorithms for each dataset Load data from and save to CSV, JSON, Numpy, and Pandas formats Who This Book Is For The

aspiring data scientist yearning to break into machine learning through mastering the underlying fundamentals that are sometimes skipped over in the rush to be productive. Some knowledge of object-oriented programming and very basic applied linear algebra will make learning easier, although anyone can benefit from this book.

Essential Statistics for Non-STEM Data Analysts - Rongpeng Li 2020-11-12

Reinforce your understanding of data science and data analysis from a statistical perspective to extract meaningful insights from your data using Python programming Key Features Work your way through the entire data analysis pipeline with statistics concerns in mind to make reasonable decisions Understand how various data science algorithms function Build a solid foundation in statistics for data science and machine learning using Python-based examples Book Description Statistics remain the backbone of modern analysis tasks, helping you to interpret the results produced by data science pipelines. This book is a detailed guide covering the math and various statistical methods required for undertaking data science tasks. The book starts by showing you how to preprocess data and inspect distributions and correlations from a statistical perspective. You'll then get to grips with the fundamentals of statistical analysis and apply its concepts to real-world datasets. As you advance, you'll find out how statistical concepts emerge from different stages of data science pipelines, understand the summary of datasets in the language of statistics, and use it to build a solid foundation for robust data products such as explanatory models and predictive models. Once you've uncovered the working mechanism of data science algorithms, you'll cover essential concepts for efficient data collection, cleaning, mining, visualization, and analysis. Finally, you'll implement statistical methods in key machine learning tasks such as classification, regression, tree-based methods, and ensemble learning. By the end of this Essential Statistics for Non-STEM Data Analysts book, you'll have learned how to build and present a self-contained, statistics-backed data product to meet your business goals. What you will learn Find out how to grab and load data into an analysis environment Perform descriptive analysis to extract meaningful summaries from data Discover probability, parameter estimation, hypothesis tests, and experiment design best practices Get to grips with resampling and bootstrapping in Python Delve into statistical tests with variance analysis, time series analysis, and A/B test examples Understand the statistics behind popular machine learning algorithms Answer questions on statistics for data scientist interviews Who this book is for This book is an entry-level guide for data science enthusiasts, data analysts, and anyone starting out in the field of data science and looking to learn the essential statistical concepts with the help of simple explanations and examples. If you're a developer or student with a non-mathematical background, you'll find this book useful. Working knowledge of the Python programming language is required.

Python for Data Analysis - Wes McKinney 2017-09-25

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

Data Science Crash Course for Beginners with Python: Fundamentals and Practices with Python - Ai Publishing 2020-08-31

Data Science Crash Course for Beginners with Python Data Science is here to stay. The tremendous growth in the volume, velocity, and variety of data has a substantial impact on every aspect of a business. While data continues to grow exponentially, accuracy remains a problem. This is where data scientists play a decisive role. A data scientist analyzes data, discovers new insights, paints a picture, and creates a vision. And a competent data scientist will provide a business with the competitive

edge it needs and address pressing business problems. Data Science Crash Course for Beginners with Python presents you with a hands-on approach to learn data science fast. How Is This Book Different? Every book by AI Publishing has been carefully crafted. This book lays equal emphasis on the theoretical sections as well as the practical aspects of data science. Each chapter provides the theoretical background behind the numerous data science techniques, and practical examples explain the working of these techniques. In the Further Reading section of each chapter, you will find the links to informative data science posts. This book presents you with the tools and packages you need to kick-start data science projects to resolve problems of practical nature. Special emphasis is laid on the main stages of a data science pipeline--data acquisition, data preparation, exploratory data analysis, data modeling and evaluation, and interpretation of the results. In the Data Science Resources section, links to data science resources, articles, interviews, and data science newsletters are provided. The author has also put together a list of contests and competitions that you can try on your own. Another added benefit of buying this book is you get instant access to all the learning material presented with this book-- PDFs, Python codes, exercises, and references--on the publisher's website. They will not cost you an extra cent. The datasets used in this book can be downloaded at runtime, or accessed via the Resources/Datasets folder. The author simplifies your learning by holding your hand through everything. The step by step description of the installation of the software you need for implementing the various data science techniques in this book is guaranteed to make your learning easier. So, right from the beginning, you can experiment with the practical aspects of data science. You'll also find the quick course on Python programming in the second and third chapters immensely helpful, especially if you are new to Python. This book gives you access to all the codes and datasets. So, access to a computer with the internet is sufficient to get started. The topics covered include: Introduction to Data Science and Decision Making Python Installation and Libraries for Data Science Review of Python for Data Science Data Acquisition Data Preparation (Preprocessing) Exploratory Data Analysis Data Modeling and Evaluation Using Machine Learning Interpretation and Reporting of Findings Data Science Projects Key Insights and Further Avenues Click the BUY button to start your Data Science journey.

Data Science from Scratch - Joel Grus 2015-04-14

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

Hands-On Data Science and Python Machine Learning - Frank Kane 2017-07-31

This book covers the fundamentals of machine learning with Python in a concise and dynamic manner. It covers data mining and large-scale machine learning using Apache Spark. About This Book Take your first steps in the world of data science by understanding the tools and techniques of data analysis Train efficient Machine Learning models in Python using the supervised and unsupervised learning methods Learn how to use Apache Spark for processing Big Data efficiently Who This Book Is For If you are a budding data scientist or a data analyst who wants to analyze and gain actionable insights from data using Python, this book is for you. Programmers with some experience in Python who want to enter the lucrative world of Data Science will also find this book to be very useful, but you don't need to be an expert Python coder or mathematician to get the most from this book. What You Will Learn Learn how to clean your data and ready it for analysis Implement the popular clustering and regression methods in Python Train efficient machine learning models using decision trees and random forests

Visualize the results of your analysis using Python's Matplotlib library Use Apache Spark's MLlib package to perform machine learning on large datasets In Detail Join Frank Kane, who worked on Amazon and IMDb's machine learning algorithms, as he guides you on your first steps into the world of data science. Hands-On Data Science and Python Machine Learning gives you the tools that you need to understand and explore the core topics in the field, and the confidence and practice to build and analyze your own machine learning models. With the help of interesting and easy-to-follow practical examples, Frank Kane explains potentially complex topics such as Bayesian methods and K-means clustering in a way that anybody can understand them. Based on Frank's successful data science course, Hands-On Data Science and Python Machine Learning empowers you to conduct data analysis and perform efficient machine learning using Python. Let Frank help you unearth the value in your data using the various data mining and data analysis techniques available in Python, and to develop efficient predictive models to predict future results. You will also learn how to perform large-scale machine learning on Big Data using Apache Spark. The book covers preparing your data for analysis, training machine learning models, and visualizing the final data analysis. Style and approach This comprehensive book is a perfect blend of theory and hands-on code examples in Python which can be used for your reference at any time.

Python for Data Science - Andrew Park 2021-02-12

☐ 55% OFF for Bookstores! NOW at \$ 36,97 instead of \$ 46.97! ☐ Did you know that according to Harvard Business Review the Data Scientist is the sexiest job of the 21st century? And for a reason! Your Customers Will Never Stop To Use This Amazing Guide! If "sexy" means having rare qualities that are much in demand, data scientists are already there. They are expensive to hire and, given the very competitive market for their services, difficult to retain. There simply aren't a lot of people with their combination of scientific background and computational and analytical skills. Data Science is all about transforming data into business value using math and algorithms. And needless to say, Python is the must-know programming language of the 21st century. If you are interested in coding and Data Science, then you must know Python to succeed in these industries! Python for Data Science is the perfect place to start learning everything you need to succeed. Contained within this guidebook are the methods, concepts, and important practical examples to help build your foundation for excelling at the discipline that is shaping the modern world. This book is perfect for programmers, software engineers, project managers and those who just want to keep up with technology. With this guide in your hands, you will: - Learn Python from scratch including the basic operations, how to install it, data structures and functions, and conditional loops - Build upon the fundamentals with advanced techniques like Object-Oriented Programming (OOP), Inheritance, and Polymorphism - Discover the importance of Data Science and how to use it in real-world situations - Learn more about Data Analysis so you can comprehend and analyze data sitting right in front of you - Increase your income by learning a new, valuable skill that only a select handful of people take the time to learn -

Mathematics for Machine Learning - Marc Peter Deisenroth 2020-04-23

The fundamental mathematical tools needed to understand machine learning include linear algebra, analytic geometry, matrix decompositions, vector calculus, optimization, probability and statistics. These topics are traditionally taught in disparate courses, making it hard for data science or computer science students, or professionals, to efficiently learn the mathematics. This self-contained textbook bridges the gap between mathematical and machine learning texts, introducing the mathematical concepts with a minimum of prerequisites. It uses these concepts to derive four central machine learning methods: linear regression, principal component analysis, Gaussian mixture models and support vector machines. For students and others with a mathematical background, these derivations provide a starting point to machine learning texts. For those learning the mathematics for the first time, the methods help build intuition and practical experience with applying mathematical concepts. Every chapter includes worked examples and exercises to test understanding. Programming tutorials are offered on the book's web site.

Principles of Data Science - Sinan Ozdemir 2016-12-16

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.

Applied Data Science with Python and Jupyter - Alex Galea 2018-10-31 Become the master player of data exploration by creating reproducible data processing pipelines, visualizations, and prediction models for your applications. Key Features Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts such as SVM, KNN classifiers, and Random Forests Discover how you can use web scraping to gather and parse your own bespoke datasets Book Description Getting started with data science doesn't have to be an uphill battle. Applied Data Science with Python and Jupyter is a step-by-step guide ideal for beginners who know a little Python and are looking for a quick, fast-paced introduction to these concepts. In this book, you'll learn every aspect of the standard data workflow process, including collecting, cleaning, investigating, visualizing, and modeling data. You'll start with the basics of Jupyter, which will be the backbone of the book. After familiarizing ourselves with its standard features, you'll look at an example of it in practice with our first analysis. In the next lesson, you dive right into predictive analytics, where multiple classification algorithms are implemented. Finally, the book ends by looking at data collection techniques. You'll see how web data can be acquired with scraping techniques and via APIs, and then briefly explore interactive visualizations. What you will learn Get up and running with the Jupyter ecosystem Identify potential areas of investigation and perform exploratory data analysis Plan a machine learning classification strategy and train classification models Use validation curves and dimensionality reduction to tune and enhance your models Scrape tabular data from web pages and transform it into Pandas DataFrames Create interactive, web-friendly visualizations to clearly communicate your findings Who this book is for Applied Data Science with Python and Jupyter is ideal for professionals with a variety of job descriptions across a 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 such as Pandas, Matplotlib, and Pandas providing you a useful head start.

Data Science Essentials in Python - Dmitry Zinoviev 2016-08-10 Go from messy, unstructured artifacts stored in SQL and NoSQL databases to a neat, well-organized dataset with this quick reference for

the busy data scientist. Understand text mining, machine learning, and network analysis; process numeric data with the NumPy and Pandas modules; describe and analyze data using statistical and network-theoretical methods; and see actual examples of data analysis at work. This one-stop solution covers the essential data science you need in Python. Data science is one of the fastest-growing disciplines in terms of academic research, student enrollment, and employment. Python, with its flexibility and scalability, is quickly overtaking the R language for data-scientific projects. Keep Python data-science concepts at your fingertips with this modular, quick reference to the tools used to acquire, clean, analyze, and store data. This one-stop solution covers essential Python, databases, network analysis, natural language processing, elements of machine learning, and visualization. Access structured and unstructured text and numeric data from local files, databases, and the Internet. Arrange, rearrange, and clean the data. Work with relational and non-relational databases, data visualization, and simple predictive analysis (regressions, clustering, and decision trees). See how typical data analysis problems are handled. And try your hand at your own solutions to a variety of medium-scale projects that are fun to work on and look good on your resume. Keep this handy quick guide at your side whether you're a student, an entry-level data science professional converting from R to Python, or a seasoned Python developer who doesn't want to memorize every function and option. What You Need: You need a decent distribution of Python 3.3 or above that includes at least NLTK, Pandas, NumPy, Matplotlib, Networkx, SciKit-Learn, and BeautifulSoup. A great distribution that meets the requirements is Anaconda, available for free from www.continuum.io. If you plan to set up your own database servers, you also need MySQL (www.mysql.com) and MongoDB (www.mongodb.com). Both packages are free and run on Windows, Linux, and Mac OS.

Introducing Data Science - Davy Cielen 2016-05-02

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 Science Introducing 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

Python Data Visualization Essentials Guide - Kallur Rahman 2021-07-30

Build your data science skills. Start data visualization Using Python. Right away. Become a good data analyst by creating quality data visualizations using Python. KEY FEATURES ● Exciting coverage on loads of Python libraries, including Matplotlib, Seaborn, Pandas, and Plotly. ● Tons of examples, illustrations, and use-cases to demonstrate visual storytelling of varied datasets. ● Covers a strong fundamental understanding of exploratory data analysis (EDA), statistical modeling, and data mining. DESCRIPTION Data visualization plays a major role in

solving data science challenges with various capabilities it offers. This book aims to equip you with a sound knowledge of Python in conjunction with the concepts you need to master to succeed as a data visualization expert. The book starts with a brief introduction to the world of data visualization and talks about why it is important, the history of visualization, and the capabilities it offers. You will learn how to do simple Python-based visualization with examples with progressive complexity of key features. The book starts with Matplotlib and explores the power of data visualization with over 50 examples. It then explores the power of data visualization using one of the popular exploratory data analysis-oriented libraries, Pandas. The book talks about statistically inclined data visualization libraries such as Seaborn. The book also teaches how we can leverage bokeh and Plotly for interactive data visualization. Each chapter is enriched and loaded with 30+ examples that will guide you in learning everything about data visualization and storytelling of mixed datasets. WHAT YOU WILL LEARN ● Learn to work with popular Python libraries and frameworks, including Seaborn, Bokeh, and Plotly. ● Practice your data visualization understanding across numerous datasets and real examples. ● Learn to visualize geospatial and time-series datasets. ● Perform correlation and EDA analysis using Pandas and Matplotlib. ● Get to know storytelling of complex and unstructured data using Bokeh and Pandas. ● Learn best practices in writing clean and short python scripts for a quicker visual summary of datasets. WHO THIS BOOK IS FOR This book is for all data analytics professionals, data scientists, and data mining hobbyists who want to be strong data visualizers by learning all the popular Python data visualization libraries. Prior working knowledge of Python is assumed. TABLE OF CONTENTS 1. Introduction to Data Visualization 2. Why Data Visualization 3. Various Data Visualization Elements and Tools 4. Using Matplotlib with Python 5. Using NumPy and Pandas for Plotting 6. Using Seaborn for Visualization 7. Using Bokeh with Python 8. Using Plotly, Folium, and Other Tools for Data Visualization 9. Hands-on Examples and Exercises, Case Studies, and Further Resources

Practical Statistics for Data Scientists - Peter Bruce 2017-05-10

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

Introduction to Data Science - Laura Igual 2017-02-26

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

Text Mining with R - Julia Silge 2017-06-12

Chapter 7. Case Study : Comparing Twitter Archives; Getting the Data and Distribution of Tweets; Word Frequencies; Comparing Word Usage; Changes in Word Use; Favorites and Retweets; Summary; Chapter 8. Case Study : Mining NASA Metadata; How Data Is Organized at NASA; Wrangling and Tidying the Data; Some Initial Simple Exploration; Word Co-occurrences and Correlations; Networks of Description and Title Words; Networks of Keywords; Calculating tf-idf for the Description

Fields; What Is tf-idf for the Description Field Words?; Connecting Description Fields to Keywords; Topic Modeling.

Fundamentals of Data Science - Samuel Burns 2019-09-17

"This book is for students or anyone, with limited or no prior programming, statistics, and data analytics knowledge. This short guide is ideal for absolute beginners, or anyone who wants to acquire a basic working knowledge of data science. It is an excellent guide if you want to learn about the principals of data science from scratch, in just a few hours. The author discussed everything that you need to know about data science. First, you are guided to learn the meaning of data science. The history of data science has been discussed to help you know how people came to realize that data is a rich source of knowledge and intelligence. The theories underlying data science have been discussed. Examples include decision and estimation theories. The author discussed the various machine learning algorithms used in data science and the various steps one has to undergo when performing data science tasks, from data collection to data presentation and visualization. The author helps you to know the various ways through which you can apply data science in your business for increased profits. A simple language has been used to ensure ease of understanding, especially for beginners." --

Foundations of Data Science - Avrim Blum 2020-01-23

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

Python Data Science Essentials - Alberto Boschetti 2016-10-28

Become an efficient data science practitioner by understanding Python's key concepts About This Book Quickly get familiar with data science using Python 3.5 Save time (and effort) with all the essential tools explained Create effective data science projects and avoid common pitfalls with the help of examples and hints dictated by experience Who This Book Is For If you are an aspiring data scientist and you have at least a working knowledge of data analysis and Python, this book will get you started in data science. Data analysts with experience of R or MATLAB will also find the book to be a comprehensive reference to enhance their data manipulation and machine learning skills. What You Will Learn Set up your data science toolbox using a Python scientific environment on Windows, Mac, and Linux Get data ready for your data science project Manipulate, fix, and explore data in order to solve data science problems Set up an experimental pipeline to test your data science hypotheses Choose the most effective and scalable learning algorithm for your data science tasks Optimize your machine learning models to get the best performance Explore and cluster graphs, taking advantage of interconnections and links in your data In Detail Fully expanded and upgraded, the second edition of Python Data Science Essentials takes you through all you need to know to succeed in data science using Python. Get modern insight into the core of Python data, including the latest versions of Jupyter notebooks, NumPy, pandas and scikit-learn. Look beyond the fundamentals with beautiful data visualizations with Seaborn and ggplot, web development with Bottle, and even the new frontiers of deep learning with Theano and TensorFlow. Dive into building your essential Python 3.5 data science toolbox, using a single-source approach that will allow to to work with Python 2.7 as well. Get to grips fast with data munging and preprocessing, and all the techniques you need to load, analyse, and process your data. Finally, get a complete overview of principal machine learning algorithms, graph analysis techniques, and all the visualization and deployment instruments that make it easier to present your results to an audience of both data science experts and business users. Style and approach The book is structured as a data science project. You will always benefit from clear code and simplified examples to help you understand the underlying mechanics and real-world datasets.

Data Science Fundamentals for Python and MongoDB - David Paper
2018-05-10

Build the foundational data science skills necessary to work with and better understand complex data science algorithms. This example-driven book provides complete Python coding examples to complement and clarify data science concepts, and enrich the learning experience. Coding examples include visualizations whenever appropriate. The book is a necessary precursor to applying and implementing machine learning algorithms. The book is self-contained. All of the math, statistics, stochastic, and programming skills required to master the content are covered. In-depth knowledge of object-oriented programming isn't required because complete examples are provided and explained. Data Science Fundamentals with Python and MongoDB is an excellent starting point for those interested in pursuing a career in data science. Like any science, the fundamentals of data science are a prerequisite to

competency. Without proficiency in mathematics, statistics, data manipulation, and coding, the path to success is "rocky" at best. The coding examples in this book are concise, accurate, and complete, and perfectly complement the data science concepts introduced. What You'll Learn Prepare for a career in data science Work with complex data structures in Python Simulate with Monte Carlo and Stochastic algorithms Apply linear algebra using vectors and matrices Utilize complex algorithms such as gradient descent and principal component analysis Wrangle, cleanse, visualize, and problem solve with data Use MongoDB and JSON to work with data Who This Book Is For The novice yearning to break into the data science world, and the enthusiast looking to enrich, deepen, and develop data science skills through mastering the underlying fundamentals that are sometimes skipped over in the rush to be productive. Some knowledge of object-oriented programming will make learning easier.