

Web Scraping With Python Successfully Scrape Data From Any Website With The Power Of Python Community Experience Distilled

Recognizing the way ways to acquire this book **web scraping with python successfully scrape data from any website with the power of python community experience distilled** is additionally useful. You have remained in right site to start getting this info. get the web scraping with python successfully scrape data from any website with the power of python community experience distilled colleague that we give here and check out the link.

You could purchase guide web scraping with python successfully scrape data from any website with the power of python community experience distilled or acquire it as soon as feasible. You could speedily download this web scraping with python successfully scrape data from any website with the power of python community experience distilled after getting deal. So, subsequent to you require the books swiftly, you can straight get it. Its consequently totally simple and thus fats, isnt it? You have to favor to in this sky

[Blueprints for Text Analytics Using Python](#) - Jens Albrecht 2020-12-04

Turning text into valuable information is essential for businesses looking to gain a competitive advantage. With recent improvements in natural language processing (NLP), users now have many options for solving complex challenges. But it's not always clear which NLP tools or libraries would work for a business's needs, or which techniques you should use and in what order. This practical book provides data scientists and developers with blueprints for best practice solutions to common tasks in text analytics and natural language processing. Authors Jens Albrecht, Sidharth Ramachandran, and Christian Winkler provide real-world case studies and detailed code examples in Python to help you get started quickly. Extract data from APIs and web pages Prepare textual data for statistical analysis and machine learning Use machine learning for classification, topic modeling, and summarization Explain AI models and classification results Explore and visualize semantic similarities with word embeddings Identify customer sentiment in product reviews Create a knowledge graph based on named entities and their relations

Research Exposed - Eszter Hargittai 2020-12-15

The era of digital communication provides endless opportunities for the collection and analysis of social data in novel ways. It also presents new and unanticipated challenges, as researchers are often inventing elements of their methodologies on the fly or studying a phenomenon or media platform for the first time. Research Exposed offers in-depth, behind-the-scenes accounts of doing empirical social science in this new paradigm. Through firsthand descriptions of innovative research projects, it shares lessons learned from over a dozen scholars' cutting-edge work. These candid accounts describe what can go wrong when pioneering new genres of research and how such difficulties can be overcome, giving both big-picture reflection and actionable advice. The chapters discuss a variety of methods, ranging from the completely novel to the use of more traditional approaches in the digital context, and cover research questions relevant to a range of disciplines, including sociology, political science, communication, information studies, and anthropology. By focusing attention on the concrete details seldom discussed in final project write-ups or traditional research guides, Research Exposed helps equip junior and senior scholars alike with essential information that is all too often left with no outlet for sharing. It offers important insights into how empirical social science research can be both innovative and rigorous when dealing with the opportunities and challenges presented by digital media.

[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

[Python Web Scraping, Second Edition](#) - Katharine Jarmul 2017-05-30

Successfully scrape data from any website with the power of Python 3.x About This Book* A hands-on guide to web scraping using Python with solutions to real-world problems* Create a number of different web scrapers in Python to extract information* This book includes practical examples on using the popular and well-maintained libraries in Python for your web scraping needs Who This Book Is For This book is aimed at developers who want to use web scraping for legitimate purposes. Prior programming experience with Python would be useful but not essential. Anyone with general knowledge of programming languages should be able to pick up the book and understand the principals involved. What You Will Learn* Extract data from web pages with simple Python programming* Build a concurrent crawler to process web pages in parallel* Follow links to crawl a website* Extract features from the HTML* Cache downloaded HTML for reuse* Compare concurrent models to determine the fastest crawler* Find out how to parse JavaScript-dependent websites* Interact with forms and sessions In Detail The Internet contains the most useful set of data ever assembled, most of which is publicly accessible for free. However, this data is not easily usable. It is embedded within the structure and style of websites and needs to be carefully extracted. Web scraping is becoming increasingly useful as a means to gather and make sense of the wealth of information available online. This book is the ultimate guide to using the latest features of Python 3.x to scrape data from websites. In the early chapters, you'll see how to extract data from static web pages. You'll learn to use caching with databases and files to save time and manage the load on servers. After covering the basics, you'll get hands-on practice building a more sophisticated crawler using browsers, crawlers, and concurrent scrapers. You'll determine when and how to scrape data from a JavaScript-dependent website using PyQt and Selenium. You'll get a better understanding of how to submit forms on complex websites protected by CAPTCHA. You'll find out how to automate these actions with Python packages such as mechanize. You'll also learn how to create class-based scrapers with Scrapy libraries and implement your learning on real websites. By the end of the book, you will have explored testing websites with scrapers, remote scraping, best practices, working with images, and many other relevant topics. Style and approach This hands-on guide is full of real-life examples and solutions starting simple and then progressively becoming more complex. Each chapter in this book introduces a problem and then provides one or more possible solutions.

Web Crawling - Christopher Olston 2010

This is a survey of the science and practice of web crawling. While at first glance web crawling may appear

to be merely an application of breadth-first-search, the truth is that there are many challenges ranging from systems concerns such as managing very large data structures, to theoretical questions such as how often to revisit evolving content sources. This survey outlines the fundamental challenges and describes the state-of-the-art models and solutions. It also highlights avenues for future work.

[Go Web Scraping Quick Start Guide](#) - Vincent Smith 2019-01-30

Web scraping is the process of extracting information from the web using various tools that perform scraping and crawling. Go is emerging as the language of choice for scraping using a variety of libraries. This book will quickly explain to you, how to scrape data data from various websites using Go libraries such as Colly and Goquery.

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

[Python Web Scraping](#) - Katharine Jarmul 2017-05-30

Successfully scrape data from any website with the power of Python 3.x About This Book A hands-on guide to web scraping using Python with solutions to real-world problems Create a number of different web scrapers in Python to extract information This book includes practical examples on using the popular and well-maintained libraries in Python for your web scraping needs Who This Book Is For This book is aimed at developers who want to use web scraping for legitimate purposes. Prior programming experience with Python would be useful but not essential. Anyone with general knowledge of programming languages should be able to pick up the book and understand the principals involved. What You Will Learn Extract data from web pages with simple Python programming Build a concurrent crawler to process web pages in parallel Follow links to crawl a website Extract features from the HTML Cache downloaded HTML for reuse Compare concurrent models to determine the fastest crawler Find out how to parse JavaScript-dependent websites Interact with forms and sessions In Detail The Internet contains the most useful set of data ever assembled, most of which is publicly accessible for free. However, this data is not easily usable. It is embedded within the structure and style of websites and needs to be carefully extracted. Web scraping is becoming increasingly useful as a means to gather and make sense of the wealth of information available online. This book is the ultimate guide to using the latest features of Python 3.x to scrape data from websites. In the early chapters, you'll see how to extract data from static web pages. You'll learn to use caching with databases and files to save time and manage the load on servers. After covering the basics, you'll get hands-on practice building a more sophisticated crawler using browsers, crawlers, and concurrent scrapers. You'll determine when and how to scrape data from a JavaScript-dependent website using PyQt and Selenium. You'll get a better understanding of how to submit forms on complex websites protected by CAPTCHA. You'll find out how to automate these actions with Python packages such as mechanize. You'll also learn how to create class-based scrapers with Scrapy libraries and implement your learning on real websites. By the end of the book, you will have explored testing websites with scrapers, remote scraping, best practices, working with images, and many other relevant topics. Style and approach This hands-on guide is full of real-life examples and solutions starting simple and then progressively becoming more complex. Each chapter in this book introduces a problem and then provides one or more possible solutions.

[Web Scraping with Python](#) - Ryan Mitchell 2015-06-15

Learn web scraping and crawling techniques to access unlimited data from any web source in any format. With this practical guide, you'll learn how to use Python scripts and web APIs to gather and process data from thousands—or even millions—of web pages at once. Ideal for programmers, security professionals, and web administrators familiar with Python, this book not only teaches basic web scraping mechanics, but also delves into more advanced topics, such as analyzing raw data or using scrapers for frontend website testing. Code samples are available to help you understand the concepts in practice. Learn how to parse complicated HTML pages Traverse multiple pages and sites Get a general overview of APIs and how they work Learn several methods for storing the data you scrape Download, read, and extract data from documents Use tools and techniques to clean badly formatted data Read and write natural languages Crawl through forms and logins Understand how to scrape JavaScript Learn image processing and text recognition

[Practical Web Scraping for Data Science](#) - Seppe vanden Broucke 2018-04-18

This book provides a complete and modern guide to web scraping, using Python as the programming language, without glossing over important details or best practices. Written with a data science audience in mind, the book explores both scraping and the larger context of web technologies in which it operates, to ensure full understanding. The authors recommend web scraping as a powerful tool for any data scientist's arsenal, as many data science projects start by obtaining an appropriate data set. Starting with a brief overview on scraping and real-life use cases, the authors explore the core concepts of HTTP, HTML, and CSS to provide a solid foundation. Along with a quick Python primer, they cover Selenium for JavaScript-heavy sites, and web crawling in detail. The book finishes with a recap of best practices and a collection of examples that bring together everything you've learned and illustrate various data science use cases. What You'll Learn Leverage well-established best practices and commonly-used Python packages Handle today's web, including JavaScript, cookies, and common web scraping mitigation techniques Understand the managerial and legal concerns regarding web scraping Who This Book is For A data science oriented audience that is probably already familiar with Python or another programming language or analytical toolkit (R, SAS, SPSS, etc). Students or instructors in university courses may also benefit. Readers unfamiliar with Python will appreciate a quick Python primer in chapter 1 to catch up with the basics and provide pointers to other guides as well.

[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.

Python Natural Language Processing - Jalaj Thanaki 2017-07-31

Leverage the power of machine learning and deep learning to extract information from text data About This Book Implement Machine Learning and Deep Learning techniques for efficient natural language processing Get started with NLTK and implement NLP in your applications with ease Understand and interpret human languages with the power of text analysis via Python Who This Book Is For This book is intended for Python developers who wish to start with natural language processing and want to make their applications smarter by implementing NLP in them. What You Will Learn Focus on Python programming paradigms, which are used to develop NLP applications Understand corpus analysis and different types of data attribute. Learn NLP using Python libraries such as NLTK, Polyglot, SpaCy, Stanford CoreNLP and so on Learn about Features Extraction and Feature selection as part of Features Engineering. Explore the advantages of vectorization in Deep Learning. Get a better understanding of the architecture of a rule-based system. Optimize and fine-tune Supervised and Unsupervised Machine Learning algorithms for NLP problems. Identify Deep Learning techniques for Natural Language Processing and Natural Language Generation problems. In Detail This book starts off by laying the foundation for Natural Language Processing and why Python is one of the best options to build an NLP-based expert system with advantages such as Community support, availability of frameworks and so on. Later it gives you a better understanding of available free forms of corpus and different types of dataset. After this, you will know how to choose a dataset for natural language processing applications and find the right NLP techniques to process sentences in datasets and understand their structure. You will also learn how to tokenize different parts of sentences and ways to analyze them. During the course of the book, you will explore the semantic as well as syntactic analysis of text. You will understand how to solve various ambiguities in processing human language and will come across various scenarios while performing text analysis. You will learn the very basics of getting the environment ready for natural language processing, move on to the initial setup, and then quickly understand sentences and language parts. You will learn the power of Machine Learning and Deep Learning to extract information from text data. By the end of the book, you will have a clear understanding of natural language processing and will have worked on multiple examples that implement NLP in the real world. Style and approach This book teaches the readers various aspects of natural language Processing using NLTK. It takes the reader from the basic to advance level in a smooth way.

JavaScript for Modern Web Development - Alok Ranjan 2020-04-18

Beginner to Expert in Web development with JavaScript: From HTML to React-Redux KEY FEATURES - Acquire web development skills to build independent applications - Understand the basics of HTML, CSS, JavaScript, React and Redux - Create build beautiful applications using HTML, CSS, JavaScript, React and Redux - Learn how to debug and unit test your applications properly to build good end products - Follow best practices to write good quality code and build performant applications DESCRIPTION This book will take you on a complete journey of learning web development, starting right with the basics. The book begins with the history of web development and JavaScript, how it has evolved over these years, and how it still keeps growing with new features. Next, you will learn the basic pillars of web development - HTML, CSS, and JavaScript. You will learn about the functional, object-oriented programming and asynchronous behaviour, and how JavaScript provides for these. Empowered with the basics, you will proceed to learn the new features of JavaScript, ES2015, and the latest ES2019. Next, you will apply your learning to build a real application to see how the Web takes shape. At the end, you will also have an introductory section on ReactJS, one of the modern frameworks for UI development and also develop a simple weather application using React. You will be introduced to Redux as the state container for React applications. This book will conclude with an introductory look at additional topics which can be taken up to become a professional and in building enterprise level applications. WHAT WILL YOU LEARN By the end of the book, you will be

building real web applications to put your knowledge to practice. This book introduces all the concepts to get started with web application development. To further excel in this field, you really need to practice by building a lot many applications, implementing your own ideas or imitating existing websites. Also remember to practice additional examples provided in the code bundle of the book to master this field. WHO THIS BOOK IS FOR This book can be used by people who are completely new to software development and want to get into front-end web development by starting from basics. This book can also be used by JavaScript users for a quick reference to the fundamentals of HTML, CSS, JS, and learn ReactJS with Redux, as well as the new features in JavaScript ES2019. Table of Contents 1. History of JS and how it has revolutionized web development 2. HTML: Creating Web Content 3. CSS: Making content beautiful 4. JavaScript Programming: Making application Interactive 5. Functional programming with JavaScript 6. Object-Oriented JavaScript 7. Asynchronous Programming 8. What's new in ES2019 JavaScript 9. Building an application with JavaScript 10. Debugging JavaScript Applications 11. Unit test automation 12. Build and Deploy an Application 13. JavaScript Best Practices 14. Introduction to React 15. Building an application with React 16. State Management in React applications 17. Debugging, Testing, and Deploying React applications 18. What is next - for becoming a pro?

The Cucumber Book - Matt Wynne 2017-02-17

Your customers want rock-solid, bug-free software that does exactly what they expect it to do. Yet they can't always articulate their ideas clearly enough for you to turn them into code. You need Cucumber: a testing, communication, and requirements tool-all rolled into one. All the code in this book is updated for Cucumber 2.4, Rails 5, and RSpec 3.5. Express your customers' wild ideas as a set of clear, executable specifications that everyone on the team can read. Feed those examples into Cucumber and let it guide your development. Build just the right code to keep your customers happy. You can use Cucumber to test almost any system or any platform. Get started by using the core features of Cucumber and working with Cucumber's Gherkin DSL to describe-in plain language-the behavior your customers want from the system. Then write Ruby code that interprets those plain-language specifications and checks them against your application. Next, consolidate the knowledge you've gained with a worked example, where you'll learn more advanced Cucumber techniques, test asynchronous systems, and test systems that use a database. Recipes highlight some of the most difficult and commonly seen situations the authors have helped teams solve. With these patterns and techniques, test Ajax-heavy web applications with Capybara and Selenium, REST web services, Ruby on Rails applications, command-line applications, legacy applications, and more. Written by the creator of Cucumber and the co-founders of Cucumber Ltd., this authoritative guide will give you and your team all the knowledge you need to start using Cucumber with confidence. What You Need: Windows, Mac OS X (with XCode) or Linux, Ruby 1.9.2 and upwards, Cucumber 2.4, Rails 5, and RSpec 3.5

Violent Python - TJ O'Connor 2012-12-28

Violent Python shows you how to move from a theoretical understanding of offensive computing concepts to a practical implementation. Instead of relying on another attacker's tools, this book will teach you to forge your own weapons using the Python programming language. This book demonstrates how to write Python scripts to automate large-scale network attacks, extract metadata, and investigate forensic artifacts. It also shows how to write code to intercept and analyze network traffic using Python, craft and spoof wireless frames to attack wireless and Bluetooth devices, and how to data-mine popular social media websites and evade modern anti-virus. Demonstrates how to write Python scripts to automate large-scale network attacks, extract metadata, and investigate forensic artifacts Write code to intercept and analyze network traffic using Python. Craft and spoof wireless frames to attack wireless and Bluetooth devices Data-mine popular social media websites and evade modern anti-virus

Python Basics - Dan Bader 2021-03-16

Make the Leap From Beginner to Intermediate in Python... Python Basics: A Practical Introduction to Python 3 Your Complete Python Curriculum-With Exercises, Interactive Quizzes, and Sample Projects What should you learn about Python in the beginning to get a strong foundation? With Python Basics, you'll not only cover the core concepts you really need to know, but you'll also learn them in the most efficient order with the help of practical exercises and interactive quizzes. You'll know enough to be dangerous with Python, fast! Who Should Read This Book If you're new to Python, you'll get a practical, step-by-step

roadmap on developing your foundational skills. You'll be introduced to each concept and language feature in a logical order. Every step in this curriculum is explained and illustrated with short, clear code samples. Our goal with this book is to educate, not to impress or intimidate. If you're familiar with some basic programming concepts, you'll get a clear and well-tested introduction to Python. This is a practical introduction to Python that jumps right into the meat and potatoes without sacrificing substance. If you have prior experience with languages like VBA, PowerShell, R, Perl, C, C++, C#, Java, or Swift the numerous exercises within each chapter will fast-track your progress. If you're a seasoned developer, you'll get a Python 3 crash course that brings you up to speed with modern Python programming. Mix and match the chapters that interest you the most and use the interactive quizzes and review exercises to check your learning progress as you go along. If you're a self-starter completely new to coding, you'll get practical and motivating examples. You'll begin by installing Python and setting up a coding environment on your computer from scratch, and then continue from there. We'll get you coding right away so that you become competent and knowledgeable enough to solve real-world problems, fast. Develop a passion for programming by solving interesting problems with Python every day! If you're looking to break into a coding or data-science career, you'll pick up the practical foundations with this book. We won't just dump a boat load of theoretical information on you so you can "sink or swim"-instead you'll learn from hands-on, practical examples one step at a time. Each concept is broken down for you so you'll always know what you can do with it in practical terms. If you're interested in teaching others "how to Python," this will be your guidebook. If you're looking to stoke the coding flame in your coworkers, kids, or relatives-use our material to teach them. All the sequencing has been done for you so you'll always know what to cover next and how to explain it. What Python Developers Say About The Book: "Go forth and learn this amazing language using this great book." - Michael Kennedy, Talk Python "The wording is casual, easy to understand, and makes the information flow well." - Thomas Wong, Pythonista "I floundered for a long time trying to teach myself. I slogged through dozens of incomplete online tutorials. I snoozed through hours of boring screencasts. I gave up on countless cruffy books from big-time publishers. And then I found Real Python. The easy-to-follow, step-by-step instructions break the big concepts down into bite-sized chunks written in plain English. The authors never forget their audience and are consistently thorough and detailed in their explanations. I'm up and running now, but I constantly refer to the material for guidance." - Jared Nielsen, Pythonista

Python for MBAs - Mattan Griffel 2021-05-04

From the ads that track us to the maps that guide us, the twenty-first century runs on code. The business world is no different. Programming has become one of the fastest-growing topics at business schools around the world. An increasing number of MBAs are choosing to pursue careers in tech. For them and other professionals, having some basic coding knowledge is a must. This book is an introduction to programming with Python for MBA students and others in business positions who need a crash course. One of the most popular programming languages, Python is used for tasks such as building and running websites, data analysis, machine learning, and natural-language processing. Drawing on years of experience providing instruction in this material at Columbia Business School as well as extensive backgrounds in technology, entrepreneurship, and consulting, Mattan Griffel and Daniel Guetta teach the basics of programming from scratch. Beginning with fundamentals such as variables, strings, lists, and functions, they build up to data analytics and practical ways to derive value from large and complex datasets. They focus on business use cases throughout, using the real-world example of a major restaurant chain to offer a concrete look at what Python can do. Written for business students with no previous coding experience and those in business roles that include coding or working with coding teams, Python for MBAs is an indispensable introduction to a versatile and powerful programming language.

Applied Data Science in Tourism - Roman Egger 2022-02-01

Access to large data sets has led to a paradigm shift in the tourism research landscape. Big data is enabling a new form of knowledge gain, while at the same time shaking the epistemological foundations and requiring new methods and analysis approaches. It allows for interdisciplinary cooperation between computer sciences and social and economic sciences, and complements the traditional research approaches. This book provides a broad basis for the practical application of data science approaches such

as machine learning, text mining, social network analysis, and many more, which are essential for interdisciplinary tourism research. Each method is presented in principle, viewed analytically, and its advantages and disadvantages are weighed up and typical fields of application are presented. The correct methodical application is presented with a "how-to" approach, together with code examples, allowing a wider reader base including researchers, practitioners, and students entering the field. The book is a very well-structured introduction to data science - not only in tourism - and its methodological foundations, accompanied by well-chosen practical cases. It underlines an important insight: data are only representations of reality, you need methodological skills and domain background to derive knowledge from them - Hannes Werthner, Vienna University of Technology Roman Egger has accomplished a difficult but necessary task: make clear how data science can practically support and foster travel and tourism research and applications. The book offers a well-taught collection of chapters giving a comprehensive and deep account of AI and data science for tourism - Francesco Ricci, Free University of Bozen-Bolzano This well-structured and easy-to-read book provides a comprehensive overview of data science in tourism. It contributes largely to the methodological repository beyond traditional methods. - Rob Law, University of Macau

Practical Machine Learning with Python - Dipanjan Sarkar 2017-12-20

Master the essential skills needed to recognize and solve complex problems with machine learning and deep learning. Using real-world examples that leverage the popular Python machine learning ecosystem, this book is your perfect companion for learning the art and science of machine learning to become a successful practitioner. The concepts, techniques, tools, frameworks, and methodologies used in this book will teach you how to think, design, build, and execute machine learning systems and projects successfully. Practical Machine Learning with Python follows a structured and comprehensive three-tiered approach packed with hands-on examples and code. Part 1 focuses on understanding machine learning concepts and tools. This includes machine learning basics with a broad overview of algorithms, techniques, concepts and applications, followed by a tour of the entire Python machine learning ecosystem. Brief guides for useful machine learning tools, libraries and frameworks are also covered. Part 2 details standard machine learning pipelines, with an emphasis on data processing analysis, feature engineering, and modeling. You will learn how to process, wrangle, summarize and visualize data in its various forms. Feature engineering and selection methodologies will be covered in detail with real-world datasets followed by model building, tuning, interpretation and deployment. Part 3 explores multiple real-world case studies spanning diverse domains and industries like retail, transportation, movies, music, marketing, computer vision and finance. For each case study, you will learn the application of various machine learning techniques and methods. The hands-on examples will help you become familiar with state-of-the-art machine learning tools and techniques and understand what algorithms are best suited for any problem. Practical Machine Learning with Python will empower you to start solving your own problems with machine learning today! What You'll Learn Execute end-to-end machine learning projects and systems Implement hands-on examples with industry standard, open source, robust machine learning tools and frameworks Review case studies depicting applications of machine learning and deep learning on diverse domains and industries Apply a wide range of machine learning models including regression, classification, and clustering. Understand and apply the latest models and methodologies from deep learning including CNNs, RNNs, LSTMs and transfer learning. Who This Book Is For IT professionals, analysts, developers, data scientists, engineers, graduate students

Python Automation Cookbook - Jaime Buelta 2020-05-29

Get a firm grip on the core processes including browser automation, web scraping, Word, Excel, and GUI automation with Python 3.8 and higher Key FeaturesAutomate integral business processes such as report generation, email marketing, and lead generationExplore automated code testing and Python's growth in data science and AI automation in three new chaptersUnderstand techniques to extract information and generate appealing graphs, and reports with MatplotlibBook Description In this updated and extended version of Python Automation Cookbook, each chapter now comprises the newest recipes and is revised to align with Python 3.8 and higher. The book includes three new chapters that focus on using Python for test automation, machine learning projects, and for working with messy data. This edition will enable you to

develop a sharp understanding of the fundamentals required to automate business processes through real-world tasks, such as developing your first web scraping application, analyzing information to generate spreadsheet reports with graphs, and communicating with automatically generated emails. Once you grasp the basics, you will acquire the practical knowledge to create stunning graphs and charts using Matplotlib, generate rich graphics with relevant information, automate marketing campaigns, build machine learning projects, and execute debugging techniques. By the end of this book, you will be proficient in identifying monotonous tasks and resolving process inefficiencies to produce superior and reliable systems. What you will learn

Learn data wrangling with Python and Pandas for your data science and AI projects
Automate tasks such as text classification, email filtering, and web scraping with Python
Use Matplotlib to generate a variety of stunning graphs, charts, and maps
Automate a range of report generation tasks, from sending SMS and email campaigns to creating templates, adding images in Word, and even encrypting PDFs
Master web scraping and web crawling of popular file formats and directories with tools like BeautifulSoup
Build cool projects such as a Telegram bot for your marketing campaign, a reader from a news RSS feed, and a machine learning model to classify emails to the correct department based on their content
Create fire-and-forget automation tasks by writing cron jobs, log files, and regexes with Python scripting

Who this book is for
Python Automation Cookbook - Second Edition is for developers, data enthusiasts or anyone who wants to automate monotonous manual tasks related to business processes such as finance, sales, and HR, among others. Working knowledge of Python is all you need to get started with this book.

Web Scraping with Python - Ryan Mitchell 2015-06-15

Learn web scraping and crawling techniques to access data from any web source in any format. Teaches basic web scraping mechanics, but also delves into more advanced topics, such as analyzing raw data or using scrapers for frontend website testing.

Automate the Boring Stuff with Python, 2nd Edition - Al Sweigart 2019-11-12

The second edition of this best-selling Python book (over 500,000 copies sold!) uses Python 3 to teach even the technically uninclined how to write programs that do in minutes what would take hours to do by hand. There is no prior programming experience required and the book is loved by liberal arts majors and geeks alike. If you've ever spent hours renaming files or updating hundreds of spreadsheet cells, you know how tedious tasks like these can be. But what if you could have your computer do them for you? In this fully revised second edition of the best-selling classic *Automate the Boring Stuff with Python*, you'll learn how to use Python to write programs that do in minutes what would take you hours to do by hand--no prior programming experience required. You'll learn the basics of Python and explore Python's rich library of modules for performing specific tasks, like scraping data off websites, reading PDF and Word documents, and automating clicking and typing tasks. The second edition of this international fan favorite includes a brand-new chapter on input validation, as well as tutorials on automating Gmail and Google Sheets, plus tips on automatically updating CSV files. You'll learn how to create programs that effortlessly perform useful feats of automation to:

- Search for text in a file or across multiple files
- Create, update, move, and rename files and folders
- Search the Web and download online content
- Update and format data in Excel spreadsheets of any size
- Split, merge, watermark, and encrypt PDFs
- Send email responses and text notifications
- Fill out online forms

Step-by-step instructions walk you through each program, and updated practice projects at the end of each chapter challenge you to improve those programs and use your newfound skills to automate similar tasks. Don't spend your time doing work a well-trained monkey could do. Even if you've never written a line of code, you can make your computer do the grunt work. Learn how in *Automate the Boring Stuff with Python, 2nd Edition*.

SQL in 10 Minutes a Day, Sams Teach Yourself - Ben Forta 2019-10-29

Sams Teach Yourself SQL in 10 Minutes offers straightforward, practical answers when you need fast results. By working through the book's 22 lessons of 10 minutes or less, you'll learn what you need to know to take advantage of the SQL language. Lessons cover IBM DB2, Microsoft SQL Server and SQL Server Express, MariaDB, MySQL, Oracle and Oracle express, PostgreSQL, and SQLite. Full-color code examples help you understand how SQL statements are structured. Tips point out shortcuts and solutions. Cautions help you avoid common pitfalls. Notes explain additional concepts, and provide additional information. 10 minutes is all you need to learn how to... Use the major SQL statements
Construct complex SQL statements

using multiple clauses and operators
Retrieve, sort, and format database contents
Pinpoint the data you need using a variety of filtering techniques
Use aggregate functions to summarize data
Join two or more related tables
Insert, update, and delete data
Create and alter database tables
Work with views, stored procedures, and more

Mining the Social Web - Matthew Russell 2011-01-21

Provides information on data analysis from a variety of social networking sites, including Facebook, Twitter, and LinkedIn.

Python Tricks - Dan Bader 2017

"I don't even feel like I've scratched the surface of what I can do with Python" With *Python Tricks: The Book* you'll discover Python's best practices and the power of beautiful & Pythonic code with simple examples and a step-by-step narrative. You'll get one step closer to mastering Python, so you can write beautiful and idiomatic code that comes to you naturally. Learning the ins and outs of Python is difficult-and with this book you'll be able to focus on the practical skills that really matter. Discover the "hidden gold" in Python's standard library and start writing clean and Pythonic code today. Who Should Read This Book: If you're wondering which lesser known parts in Python you should know about, you'll get a roadmap with this book. Discover cool (yet practical!) Python tricks and blow your coworkers' minds in your next code review. If you've got experience with legacy versions of Python, the book will get you up to speed with modern patterns and features introduced in Python 3 and backported to Python 2. If you've worked with other programming languages and you want to get up to speed with Python, you'll pick up the idioms and practical tips you need to become a confident and effective Pythonista. If you want to make Python your own and learn how to write clean and Pythonic code, you'll discover best practices and little-known tricks to round out your knowledge. What Python Developers Say About The Book: "I kept thinking that I wished I had access to a book like this when I started learning Python many years ago." - Mariatta Wijaya, Python Core Developer "This book makes you write better Python code!" - Bob Belderbos, Software Developer at Oracle "Far from being just a shallow collection of snippets, this book will leave the attentive reader with a deeper understanding of the inner workings of Python as well as an appreciation for its beauty." - Ben Felder, Pythonista "It's like having a seasoned tutor explaining, well, tricks!" - Daniel Meyer, Sr. Desktop Administrator at Tesla Inc.

Data Wrangling with Python - Jacqueline Kazil 2016-02-04

How do you take your data analysis skills beyond Excel to the next level? By learning just enough Python to get stuff done. This hands-on guide shows non-programmers like you how to process information that's initially too messy or difficult to access. You don't need to know a thing about the Python programming language to get started. Through various step-by-step exercises, you'll learn how to acquire, clean, analyze, and present data efficiently. You'll also discover how to automate your data process, schedule file- editing and clean-up tasks, process larger datasets, and create compelling stories with data you obtain. Quickly learn basic Python syntax, data types, and language concepts
Work with both machine-readable and human-consumable data
Scrape websites and APIs to find a bounty of useful information
Clean and format data to eliminate duplicates and errors in your datasets
Learn when to standardize data and when to test and script data cleanup
Explore and analyze your datasets with new Python libraries and techniques
Use Python solutions to automate your entire data-wrangling process

Encyclopedia of Big Data - Laurie A. Schintler 2022-02-23

This encyclopedia will be an essential resource for our times, reflecting the fact that we currently are living in an expanding data-driven world. Technological advancements and other related trends are contributing to the production of an astoundingly large and exponentially increasing collection of data and information, referred to in popular vernacular as "Big Data." Social media and crowdsourcing platforms and various applications — "apps" — are producing reams of information from the instantaneous transactions and input of millions and millions of people around the globe. The Internet-of-Things (IoT), which is expected to comprise tens of billions of objects by the end of this decade, is actively sensing real-time intelligence on nearly every aspect of our lives and environment. The Global Positioning System (GPS) and other location-aware technologies are producing data that is specific down to particular latitude and longitude coordinates and seconds of the day. Large-scale instruments, such as the Large Hadron Collider (LHC), are

collecting massive amounts of data on our planet and even distant corners of the visible universe. Digitization is being used to convert large collections of documents from print to digital format, giving rise to large archives of unstructured data. Innovations in technology, in the areas of Cloud and molecular computing, Artificial Intelligence/Machine Learning, and Natural Language Processing (NLP), to name only a few, also are greatly expanding our capacity to store, manage, and process Big Data. In this context, the Encyclopedia of Big Data is being offered in recognition of a world that is rapidly moving from gigabytes to terabytes to petabytes and beyond. While indeed large data sets have long been around and in use in a variety of fields, the era of Big Data in which we now live departs from the past in a number of key respects and with this departure comes a fresh set of challenges and opportunities that cut across and affect multiple sectors and disciplines, and the public at large. With expanded analytical capacities at hand, Big Data is now being used for scientific inquiry and experimentation in nearly every (if not all) disciplines, from the social sciences to the humanities to the natural sciences, and more. Moreover, the use of Big Data has been well established beyond the Ivory Tower. In today's economy, businesses simply cannot be competitive without engaging Big Data in one way or another in support of operations, management, planning, or simply basic hiring decisions. In all levels of government, Big Data is being used to engage citizens and to guide policy making in pursuit of the interests of the public and society in general. Moreover, the changing nature of Big Data also raises new issues and concerns related to, for example, privacy, liability, security, access, and even the veracity of the data itself. Given the complex issues attending Big Data, there is a real need for a reference book that covers the subject from a multi-disciplinary, cross-sectoral, comprehensive, and international perspective. The Encyclopedia of Big Data will address this need and will be the first of such reference books to do so. Featuring some 500 entries, from "Access" to "Zillow," the Encyclopedia will serve as a fundamental resource for researchers and students, for decision makers and leaders, and for business analysts and purveyors. Developed for those in academia, industry, and government, and others with a general interest in Big Data, the encyclopedia will be aimed especially at those involved in its collection, analysis, and use. Ultimately, the Encyclopedia of Big Data will provide a common platform and language covering the breadth and depth of the topic for different segments, sectors, and disciplines.

Python Social Media Analytics - Siddhartha Chatterjee 2017-07-28

Leverage the power of Python to collect, process, and mine deep insights from social media data About This Book Acquire data from various social media platforms such as Facebook, Twitter, YouTube, GitHub, and more Analyze and extract actionable insights from your social data using various Python tools A highly practical guide to conducting efficient social media analytics at scale Who This Book Is For If you are a programmer or a data analyst familiar with the Python programming language and want to perform analyses of your social data to acquire valuable business insights, this book is for you. The book does not assume any prior knowledge of any data analysis tool or process. What You Will Learn Understand the basics of social media mining Use PyMongo to clean, store, and access data in MongoDB Understand user reactions and emotion detection on Facebook Perform Twitter sentiment analysis and entity recognition using Python Analyze video and campaign performance on YouTube Mine popular trends on GitHub and predict the next big technology Extract conversational topics on public internet forums Analyze user interests on Pinterest Perform large-scale social media analytics on the cloud In Detail Social Media platforms such as Facebook, Twitter, Forums, Pinterest, and YouTube have become part of everyday life in a big way. However, these complex and noisy data streams pose a potent challenge to everyone when it comes to harnessing them properly and benefiting from them. This book will introduce you to the concept of social media analytics, and how you can leverage its capabilities to empower your business. Right from acquiring data from various social networking sources such as Twitter, Facebook, YouTube, Pinterest, and social forums, you will see how to clean data and make it ready for analytical operations using various Python APIs. This book explains how to structure the clean data obtained and store in MongoDB using PyMongo. You will also perform web scraping and visualize data using Scrapy and BeautifulSoup. Finally, you will be introduced to different techniques to perform analytics at scale for your social data on the cloud, using Python and Spark. By the end of this book, you will be able to utilize the power of Python to gain valuable insights from social media data and use them to enhance your business processes. Style and

approach This book follows a step-by-step approach to teach readers the concepts of social media analytics using the Python programming language. To explain various data analysis processes, real-world datasets are used wherever required.

Getting Structured Data from the Internet - Jay M. Patel 2020-12-13

Utilize web scraping at scale to quickly get unlimited amounts of free data available on the web into a structured format. This book teaches you to use Python scripts to crawl through websites at scale and scrape data from HTML and JavaScript-enabled pages and convert it into structured data formats such as CSV, Excel, JSON, or load it into a SQL database of your choice. This book goes beyond the basics of web scraping and covers advanced topics such as natural language processing (NLP) and text analytics to extract names of people, places, email addresses, contact details, etc., from a page at production scale using distributed big data techniques on an Amazon Web Services (AWS)-based cloud infrastructure. It book covers developing a robust data processing and ingestion pipeline on the Common Crawl corpus, containing petabytes of data publicly available and a web crawl data set available on AWS's registry of open data. Getting Structured Data from the Internet also includes a step-by-step tutorial on deploying your own crawlers using a production web scraping framework (such as Scrapy) and dealing with real-world issues (such as breaking Captcha, proxy IP rotation, and more). Code used in the book is provided to help you understand the concepts in practice and write your own web crawler to power your business ideas. What You Will Learn Understand web scraping, its applications/uses, and how to avoid web scraping by hitting publicly available rest API endpoints to directly get data Develop a web scraper and crawler from scratch using lxml and BeautifulSoup library, and learn about scraping from JavaScript-enabled pages using Selenium Use AWS-based cloud computing with EC2, S3, Athena, SQS, and SNS to analyze, extract, and store useful insights from crawled pages Use SQL language on PostgreSQL running on Amazon Relational Database Service (RDS) and SQLite using SQLAlchemy Review sci-kit learn, Gensim, and spaCy to perform NLP tasks on scraped web pages such as name entity recognition, topic clustering (Kmeans, Agglomerative Clustering), topic modeling (LDA, NMF, LSI), topic classification (naive Bayes, Gradient Boosting Classifier) and text similarity (cosine distance-based nearest neighbors) Handle web archival file formats and explore Common Crawl open data on AWS Illustrate practical applications for web crawl data by building a similar website tool and a technology profiler similar to builtwith.com Write scripts to create a backlinks database on a web scale similar to Ahrefs.com, Moz.com, Majestic.com, etc., for search engine optimization (SEO), competitor research, and determining website domain authority and ranking Use web crawl data to build a news sentiment analysis system or alternative financial analysis covering stock market trading signals Write a production-ready crawler in Python using Scrapy framework and deal with practical workarounds for Captchas, IP rotation, and more Who This Book Is For Primary audience: data analysts and scientists with little to no exposure to real-world data processing challenges, secondary: experienced software developers doing web-heavy data processing who need a primer, tertiary: business owners and startup founders who need to know more about implementation to better direct their technical team

Learning Scrapy - Dimitrios Kouzis-Loukas 2016-01-30

Learn the art of efficient web scraping and crawling with Python About This Book Extract data from any source to perform real time analytics. Full of techniques and examples to help you crawl websites and extract data within hours. A hands-on guide to web scraping and crawling with real-life problems and solutions Who This Book Is For If you are a software developer, data scientist, NLP or machine-learning enthusiast or just need to migrate your company's wiki from a legacy platform, then this book is for you. It is perfect for someone , who needs instant access to large amounts of semi-structured data effortlessly. What You Will Learn Understand HTML pages and write XPath to extract the data you need Write Scrapy spiders with simple Python and do web crawls Push your data into any database, search engine or analytics system Configure your spider to download files, images and use proxies Create efficient pipelines that shape data in precisely the form you want Use Twisted Asynchronous API to process hundreds of items concurrently Make your crawler super-fast by learning how to tune Scrapy's performance Perform large scale distributed crawls with scrapyd and scrapinghub In Detail This book covers the long awaited Scrapy v 1.0 that empowers you to extract useful data from virtually any source with very little effort. It starts off by explaining the fundamentals of Scrapy framework, followed by a thorough description of how to extract

data from any source, clean it up, shape it as per your requirement using Python and 3rd party APIs. Next you will be familiarised with the process of storing the scrapped data in databases as well as search engines and performing real time analytics on them with Spark Streaming. By the end of this book, you will perfect the art of scarping data for your applications with ease Style and approach It is a hands on guide, with first few chapters written as a tutorial, aiming to motivate you and get you started quickly. As the book progresses, more advanced features are explained with real world examples that can be referred while developing your own web applications.

[Advances in Computational and Bio-Engineering](#) - S. Jyothi 2020-10-23

This book gathers state-of-the-art research in computational engineering and bioengineering to facilitate knowledge exchange between various scientific communities. Computational engineering (CE) is a relatively new discipline that addresses the development and application of computational models and simulations often coupled with high-performance computing to solve complex physical problems arising in engineering analysis and design in the context of natural phenomena. Bioengineering (BE) is an important aspect of computational biology, which aims to develop and use efficient algorithms, data structures, and visualization and communication tools to model biological systems. Today, engineering approaches are essential for biologists, enabling them to analyse complex physiological processes, as well as for the pharmaceutical industry to support drug discovery and development programmes.

[The Hitchhiker's Guide to Python](#) - Kenneth Reitz 2016-08-30

The Hitchhiker's Guide to Python takes the journeyman Pythonista to true expertise. More than any other language, Python was created with the philosophy of simplicity and parsimony. Now 25 years old, Python has become the primary or secondary language (after SQL) for many business users. With popularity comes diversity—and possibly dilution. This guide, collaboratively written by over a hundred members of the Python community, describes best practices currently used by package and application developers. Unlike other books for this audience, The Hitchhiker's Guide is light on reusable code and heavier on design philosophy, directing the reader to excellent sources that already exist.

[Python for Excel](#) - Felix Zumstein 2021-03-04

While Excel remains ubiquitous in the business world, recent Microsoft feedback forums are full of requests to include Python as an Excel scripting language. In fact, it's the top feature requested. What makes this combination so compelling? In this hands-on guide, Felix Zumstein--creator of xlwings, a popular open source package for automating Excel with Python--shows experienced Excel users how to integrate these two worlds efficiently. Excel has added quite a few new capabilities over the past couple of years, but its automation language, VBA, stopped evolving a long time ago. Many Excel power users have already adopted Python for daily automation tasks. This guide gets you started. Use Python without extensive programming knowledge Get started with modern tools, including Jupyter notebooks and Visual Studio code Use pandas to acquire, clean, and analyze data and replace typical Excel calculations Automate tedious tasks like consolidation of Excel workbooks and production of Excel reports Use xlwings to build interactive Excel tools that use Python as a calculation engine Connect Excel to databases and CSV files and fetch data from the internet using Python code Use Python as a single tool to replace VBA, Power Query, and Power Pivot

[Python Web Scraping Cookbook](#) - Michael Heydt 2018-02-09

Untangle your web scraping complexities and access web data with ease using Python scripts Key Features Hands-on recipes for advancing your web scraping skills to expert level One-stop solution guide to address complex and challenging web scraping tasks using Python Understand web page structures and collect data from a website with ease Book Description Python Web Scraping Cookbook is a solution-focused book that will teach you techniques to develop high-performance Scrapers, and deal with cookies, hidden form fields, Ajax-based sites and proxies. You'll explore a number of real-world scenarios where every part of the development or product life cycle will be fully covered. You will not only develop the skills to design reliable, high-performing data flows, but also deploy your codebase to Amazon Web Services (AWS). If you are involved in software engineering, product development, or data mining or in building data-driven products, you will find this book useful as each recipe has a clear purpose and objective. Right from extracting data from websites to writing a sophisticated web crawler, the book's independent recipes will

be extremely helpful while on the job. This book covers Python libraries, requests, and BeautifulSoup. You will learn about crawling, web spidering, working with AJAX websites, and paginated items. You will also understand to tackle problems such as 403 errors, working with proxy, scraping images, and LXML. By the end of this book, you will be able to scrape websites more efficiently and deploy and operate your scraper in the cloud. What you will learn Use a variety of tools to scrape any website and data, including Scrapy and Selenium Master expression languages, such as XPath and CSS, and regular expressions to extract web data Deal with scraping traps such as hidden form fields, throttling, pagination, and different status codes Build robust scraping pipelines with SQS and RabbitMQ Scrape assets like image media and learn what to do when Scraper fails to run Explore ETL techniques of building a customized crawler, parser, and convert structured and unstructured data from websites Deploy and run your scraper as a service in AWS Elastic Container Service Who this book is for This book is ideal for Python programmers, web administrators, security professionals, and anyone who wants to perform web analytics. Familiarity with Python and basic understanding of web scraping will be useful to make the best of this book.

[Hands-On Web Scraping with Python](#) - Anish Chapagain 2019-07-15

Collect and scrape different complexities of data from the modern Web using the latest tools, best practices, and techniques Key Features Learn various scraping techniques using a range of Python libraries such as Scrapy and Beautiful Soup Build scrapers and crawlers to extract relevant information from the web Automate web scraping operations to bridge the accuracy gap and ease complex business needs Book Description Web scraping is an essential technique used in many organizations to scrape valuable data from web pages. This book will enable you to delve deeply into web scraping techniques and methodologies. This book will introduce you to the fundamental concepts of web scraping techniques and how they can be applied to multiple sets of web pages. We'll use powerful libraries from the Python ecosystem—such as Scrapy, lxml, pyquery, bs4, and others—to carry out web scraping operations. We will take an in-depth look at essential tasks to carry out simple to intermediate scraping operations such as identifying information from web pages, using patterns or attributes to retrieve information, and others. This book adopts a practical approach to web scraping concepts and tools, guiding you through a series of use cases and showing you how to use the best tools and techniques to efficiently scrape web pages. This book also covers the use of other popular web scraping tools, such as Selenium, Regex, and web-based APIs. By the end of this book, you will have learned how to efficiently scrape the web using different techniques with Python and other popular tools. What you will learn Analyze data and Information from web pages Learn how to use browser-based developer tools from the scraping perspective Use XPath and CSS selectors to identify and explore markup elements Learn to handle and manage cookies Explore advanced concepts in handling HTML forms and processing logins Optimize web securities, data storage, and API use to scrape data Use Regex with Python to extract data Deal with complex web entities by using Selenium to find and extract data Who this book is for This book is for Python programmers, data analysts, web scraping newbies, and anyone who wants to learn how to perform web scraping from scratch. If you want to begin your journey in applying web scraping techniques to a range of web pages, then this book is what you need! A working knowledge of the Python programming language is expected.

[Practical Docker with Python](#) - Sathyajith Bhat 2018-07-26

Learn the key differences between containers and virtual machines. Adopting a project based approach, this book introduces you to a simple Python application to be developed and containerized with Docker. After an introduction to Containers and Docker you'll be guided through Docker installation and configuration. You'll also learn basic functions and commands used in Docker by running a simple container using Docker commands. The book then moves on to developing a Python based Messaging Bot using required libraries and virtual environment where you'll add Docker Volumes to your project, ensuring your container data is safe. You'll create a database container and link your project to it and finally, bring up the Bot-associated database all at once with Docker Compose. What You'll Learn Build, run, and distribute Docker containers Develop a Python App and containerize it Use Dockerfile to run the Python App Define and run multi-container applications with Docker Compose Work with persisting data generated by and used by Docker containers Who This Book Is For Intermediate developers/DevOps practitioners who are looking to improve their build and release workflow by containerizing applications

Web Scraping for Data Science with Python - Seppe vanden Broucke 2017-11-30

Get Started with Web Scraping using Python! Congratulations! By picking up this book, you've set the first steps into the exciting world of web scraping. For those who are not familiar with programming or the deeper workings of the web, web scraping often looks like a black art: the ability to write a program that sets off on its own to explore the Internet and collect data is seen as a magical and exciting ability to possess. In this book, we set out to provide a concise and modern guide to web scraping, using Python as our programming language, without glossing over important details or best practices. In addition, this book is written with a data science audience in mind. We're data scientists ourselves, and have very often found web scraping to be a powerful tool to have in your arsenal, as many data science projects start with the first step of obtaining an appropriate data set, so why not utilize the treasure trove of information the web provides. As such, we've strived to offer a guide that: Is concise and to the point, whilst also being thorough Is geared towards data scientists: we'll show you how web scraping fits into the data science workflow Takes a "code first" approach to get you up to speed quickly without too much boilerplate text Is modern by using well-established best practices and Python packages only Shows how to handle the web of today, including JavaScript, cookies, and common web scraping mitigation techniques Includes a thorough managerial and legal discussion regarding web scraping Provides lots of pointers for further reading and learning Includes many larger, fully worked out examples Chapter Overview Nine chapters are included in this book. In Chapter 1, we provide a brief overview on web scraping and real-life use cases and make sure your Python environment is set up correctly. In Chapter 2, you'll learn the basics regarding HTTP, the core piece of technology behind the web, and the requests Python library. In Chapter 3, we start working with HTML and CSS sites, using the BeautifulSoup library. Chapter 4 returns to HTTP, exploring it more detail. Chapter 5 introduces the Selenium library, which you'll use to scrape JavaScript-heavy websites. Chapter 6 explains web crawling in detail. In Chapter 7, an in-depth discussion regarding managerial and legal concerns is provided. Chapter 8 recaps best practices and provides pointers to other tools. Chapter 9 includes fourteen, fully worked out web scraping examples bringing everything you've learned together, and illustrates various interesting data science oriented use cases.

Website Scraping with Python - Gábor László Hajba 2018

Offering road-tested techniques for website scraping and solutions to common issues developers may face, this concise and focused book provides tips and tweaking guidance for the popular scraping tools BeautifulSoup and Scrapy. --

Python for Everybody - Charles R. Severance 2016-04-09

Python for Everybody is designed to introduce students to programming and software development through the lens of exploring data. You can think of the Python programming language as your tool to solve data

problems that are beyond the capability of a spreadsheet. Python is an easy to use and easy to learn programming language that is freely available on Macintosh, Windows, or Linux computers. So once you learn Python you can use it for the rest of your career without needing to purchase any software. This book uses the Python 3 language. The earlier Python 2 version of this book is titled "Python for Informatics: Exploring Information". There are free downloadable electronic copies of this book in various formats and supporting materials for the book at www.pythonlearn.com. The course materials are available to you under a Creative Commons License so you can adapt them to teach your own Python course.

Web Scraping with Python - Richard Lawson 2015-10-28

Successfully scrape data from any website with the power of Python About This Book A hands-on guide to web scraping with real-life problems and solutions Techniques to download and extract data from complex websites Create a number of different web scrapers to extract information Who This Book Is For This book is aimed at developers who want to use web scraping for legitimate purposes. Prior programming experience with Python would be useful but not essential. Anyone with general knowledge of programming languages should be able to pick up the book and understand the principals involved. What You Will Learn Extract data from web pages with simple Python programming Build a threaded crawler to process web pages in parallel Follow links to crawl a website Download cache to reduce bandwidth Use multiple threads and processes to scrape faster Learn how to parse JavaScript-dependent websites Interact with forms and sessions Solve CAPTCHAs on protected web pages Discover how to track the state of a crawl In Detail The Internet contains the most useful set of data ever assembled, largely publicly accessible for free. However, this data is not easily reusable. It is embedded within the structure and style of websites and needs to be carefully extracted to be useful. Web scraping is becoming increasingly useful as a means to easily gather and make sense of the plethora of information available online. Using a simple language like Python, you can crawl the information out of complex websites using simple programming. This book is the ultimate guide to using Python to scrape data from websites. In the early chapters it covers how to extract data from static web pages and how to use caching to manage the load on servers. After the basics we'll get our hands dirty with building a more sophisticated crawler with threads and more advanced topics. Learn step-by-step how to use Ajax URLs, employ the Firebug extension for monitoring, and indirectly scrape data. Discover more scraping nitty-gritties such as using the browser renderer, managing cookies, how to submit forms to extract data from complex websites protected by CAPTCHA, and so on. The book wraps up with how to create high-level scrapers with Scrapy libraries and implement what has been learned to real websites. Style and approach This book is a hands-on guide with real-life examples and solutions starting simple and then progressively becoming more complex. Each chapter in this book introduces a problem and then provides one or more possible solutions.