

Web Scraping With Python, 2e

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.

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 Features Automate integral business processes such as report generation, email marketing, and lead generation Explore automated code testing and Python's growth in data science and AI automation in three new chapters Understand techniques to extract information and generate appealing graphs, and reports with Matplotlib Book 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 code projects such as a Telegram bot for your marketing campaign, a reader from a news RSS feed, and an 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.

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 scrapy and scrapyhub 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 scraping 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.

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

Black Hat Python, 2nd Edition

Perform advanced scraping operations using various Python libraries and tools such as Selenium, Regex, and others

Introducing Python

Go Web Scraping Quick Start Guide

Scrape, Clean, Explore & Transform Your Data

Fluent Python

This Book Includes : Python Basics for Beginners + Python Automation Techniques And Web Scraping + Python For Data Science And Machine Learning

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.

Perform advanced data manipulation tasks using pandas and become an expert data analyst. Key Features Manipulate and analyze your data expertly using the power of pandas Work with missing data and time series data and become a true pandas expert Includes expert tips and techniques on making your data analysis tasks easier Book Description pandas is a popular Python library used by data scientists and analysts worldwide to manipulate and analyze their data. This book presents useful data manipulation techniques in pandas to perform complex data analysis in various domains. An update to our highly successful previous edition with new features, examples, updated code, and more, this book is an in-depth guide to get the most out of pandas for data analysis. Designed for both intermediate users as well as seasoned practitioners, you will learn advanced data manipulation techniques, such as multi-indexing, modifying data structures, and sampling your data, which allow for powerful analysis and help you gain accurate insights from it. With the help of this book, you will apply pandas to different domains, such as Bayesian statistics, predictive analytics, and time series analysis using an example-based approach. And not just that; you will also learn how to prepare powerful, interactive business reports in pandas using the Jupyter notebook. By the end of this book, you will learn how to perform efficient data analysis using pandas on complex data, and become an expert data analyst or data scientist in the process. What you will learn Speed up your data analysis by importing data into pandas Keep relevant data points by selecting subsets of your data Create a high-quality dataset by cleaning data and fixing missing values Compute actionable analytics with grouping and aggregation in pandas Master time series data analysis in pandas Make powerful reports in pandas using Jupyter notebooks Who this book is for This book is for data scientists, analysts and Python developers who wish to explore advanced data analysis and scientific computing techniques using pandas. Some fundamental understanding of Python programming and familiarity with the basic data analysis concepts is all you need to get started with this book.

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 BeautifulSoup 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.

Python is an integrated, object-orientated development language for use in computer programming. This text is split into distinct sections, each concentrating on a core angle of the language. The book also contains sections for Web and application development, the two most popular uses for Python. It is designed to teach a programmer how to use Python by explaining the mechanics of Python. The appendixes offer a quick guide to the main features of the Python language, as well as additional guides to non-essential systems such as the IDLE development environment and general guidelines for migrating from another language.

A Guide to Developing Internet Agents with PHP/CURL

Techniques and tools to crawl and scrape data from websites

Python Programming for Hackers and Pentesters

Build enterprise-grade, scalable Python web applications, 2nd Edition

Techniques for ethical hacking with Python, 2nd Edition

Hands-On Web Scraping with Python

Python is fast becoming the programming language of choice for hackers, reverse engineers, and software testers because it's easy to write quickly, and it has the low-level support and libraries that make hackers happy. But until now, there has been no real manual on how to use Python for a variety of hacking tasks. You had to dig through forum posts and man pages, endlessly tweaking your own code to get everything working. Not anymore. Gray Hat Python explains the concepts behind hacking tools and techniques like debuggers, trojans, fuzzers, and emulators. But author Justin Seltz goes beyond theory, showing you how to harness existing Python-based security tools—and how to build your own when the pre-built ones won't cut it. You'll learn how to –Automate tedious reversing and security tasks –Design and program your own debugger –Learn how to fuzz Windows drivers and create powerful fuzzers from scratch –Have fun with code and library injection, soft and hard hooking techniques, and other software trickery –Sniff secure traffic out of an encrypted web browser session –Use PyDBG, Immunity Debugger, Sulley, IDAPython, PyEMU, and more The world's best hackers are using Python to do their handiwork. Shouldn't you?

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.

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

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.

High Performance Python

Data Wrangling with Pandas, NumPy, and IPython

Implement the power of Go to scrape and crawl data from the web

The Hitchhiker's Guide to Python

Clear, Concise, and Effective Programming

Webbots, Spiders, and Screen Scrapers, 2nd Edition

Mining Social Media

Web Scraping with Python Collecting More Data from the Modern Web

Provides information on ways to automate online tasks using webbots and spiders, covering such topics as parsing data from Web pages, managing cookies, sending and receiving email, and decoding encrypted files.

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 effectively. 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

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 of 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.

R Web Scraping Quick Start Guide

Python Projects for Beginners

Mining the Social Web

Python for Data Analysis

Mastering pandas

Python Penetration Testing Essentials

Python Crash Course, 2nd Edition

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 of 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.

Using Python to scrape websites 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 BeautifulSoupSoup. 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 or Scrapy assets like image media and learn what to do when Scrapy 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.

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

Fully-updated for Python 3, the second edition of this worldwide bestseller (over 100,000 copies sold) explores the stealthier side of programming and brings you all new strategies for your hacking projects. When it comes to creating powerful and effective hacking tools, Python is the language of choice for most security analysts. In Black Hat Python, 2nd Edition, you'll explore the darker side of Python's capabilities—writing network sniffers, stealing email credentials, brute forcing directories, crafting mutation fuzzers, infecting virtual machines, creating stealthy trojans, and more. The second edition of this bestselling hacking book contains code updated for the latest version of Python 3, as well as new techniques that reflect current industry best practices. You'll also find expanded explanations of Python libraries such as ctypes, struct, lxml, and BeautifulSoup, and dig deeper into strategies, from splitting bits to leveraging computer-vision libraries, that you can apply to future hacking projects. You'll learn how to:

- Create a trojan command-and-control using GitHub
- Detect sandboxing and automate common malware tasks, like keylogging and screenshots
- Escalate Windows privileges with creative process control
- Use offensive memory forensics tricks to retrieve password hashes and inject shellcode into a virtual machine
- Extend the popular Burp Suite web-hacking tool
- Abuse Windows COM automation to perform a man-in-the-browser attack

Exfiltrate data from a network most sneakily When it comes to offensive security, your ability to create powerful tools on the fly is indispensable. Learn how with the second edition of Black Hat Python. New to this edition: All Python code has been updated to cover Python 3 and includes updated libraries used in current Python applications. Additionally, there are more in-depth explanations of the code and the programming techniques have been updated to current, common tactics. Examples of new material that you'll learn include how to sniff network traffic, evade anti-virus software, brute-force web applications, and set up a command-and-control (C2) system using GitHub.

Python for Everybody

The Data Wrangling Workshop

Finding Stories in Internet Data

Python for Excel

A Ten-Week Bootcamp Approach to Python Programming

Developing Web Applications with Python

Automate the Boring Stuff with Python, 2nd Edition

This book gives you the skills you need to use Python for penetration testing, with the help of detailed code examples. This book has been updated for Python 3.6.3 and Kali Linux 2018.1. Key Features Detect and avoid various attack types that put the privacy of a system at risk Leverage Python to build efficient code and eventually build a robust environment Learn about securing wireless applications and information gathering on a web server Book Description This book gives you the skills you need to use Python for penetration testing (pentesting), with the help of detailed code examples. We start by exploring the basics of networking with Python and then proceed to network hacking. Then, you will delve into exploring Python libraries to perform various types of pentesting and ethical hacking techniques. Next, we delve into hacking the application layer, where we start by gathering information from a website. We then move on to concepts related to website hacking—such as parameter tampering, DDoS, XSS, and SQL injection. By reading this book, you will learn different techniques and methodologies that will familiarize you with Python pentesting techniques, how to protect yourself, and how to create automated programs to find the admin console, SQL injection, and XSS attacks. What you will learn The basics of network penetration testing and how to sniff network traffic for attack and torrent detection Web server footprinting and web application attacks, including the XSS and SQL injection attack Wireless frames and how to obtain information such as SSID, BSSID, and the channel number from a wireless frame using a Python script The importance of web server signatures, email gathering, and why knowing the server signature is the first step in hacking Who this book is for If you are a Python programmer, a security researcher, or an ethical hacker and are interested in penetration testing with the help of Python, then this book is for you. Even if you are new to the field of ethical hacking, this book can help you find the vulnerabilities in your system so that you are ready to tackle any kind of attack or intrusion.

Are you thinking about learning how to use the Python programming language? Thinking about getting started on a programming career? Are you thinking about learning data science? This book is for you!

Learn how to turn raw data into rich, interactive web visualizations with the powerful combination of Python and JavaScript. With this hands-on guide, author Kyrán Dale teaches you how build a basic dataviz toolchain with best-of-breed Python and JavaScript libraries—including Scrapy, Matplotlib, Pandas, Flask, and D3—for crafting engaging, browser-based visualizations. As a working example, throughout the book Dale walks you through transforming Wikipedia's table-based list of Nobel Prize winners into an interactive visualization. You'll examine steps along the entire toolchain, from scraping, cleaning, exploring, and delivering data to building the visualization with JavaScript's D3 library. If you're ready to create your own web-based data visualizations—and know either Python or JavaScript—this is the book for you. Learn how to manipulate data with Python Understand the commonalities between Python and JavaScript Extract information from websites by using Python's web-scraping tools, BeautifulSoup and Scrapy Clean and explore data with Python's Pandas, Matplotlib, and Numpy libraries Serve data and create RESTful web APIs with Python's Flask framework Create engaging, interactive web visualizations with JavaScript's D3 library 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.

Learning Scrapy

75 Python automation ideas for web scraping, data wrangling, and processing Excel, reports, emails, and more, 2nd Edition

Python Programming for Hackers and Reverse Engineers

Python

Modern Computing in Simple Packages

The Complete Reference

Best Practices and Examples with Python

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. 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.

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.

Learn to build modern, secure, highly available web MVC applications and APIs using Python's Flask framework. Key Features Create production-ready MVC and REST API with the dynamic features of Flask Utilize the various extensions like Flask-JWT and Flask-SQLAlchemy to develop powerful applications Deploy your flask application on real-world platforms like AWS and Heroku on VM's or Docker containers Book Description Flask is a popular Python framework known for its lightweight and modular design. Mastering Flask Web Development will take you on a complete tour of the Flask environment and teach you how to build a production-ready application. You'll begin by learning about the installation of Flask and basic concepts such as MVC and accessing a database using an ORM. You will learn how to structure your application so that it can scale to any size with the help of Flask Blueprints. You'll then learn how to use Jinja2 templates with a high level of expertise. You will also learn how to develop with SQL or NoSQL databases, and how to develop REST APIs and JWT authentication. Next, you'll move on to build role-based access security and authentication using LDAP, OAuth, OpenID, and database. Also learn how to create asynchronous tasks that can scale to any load using Celery and RabbitMQ or Redis. You will also be introduced to a wide range of Flask extensions to leverage technologies such as cache, localization, and debugging. You will learn how to build your own Flask extensions, how to write tests, and how to get test coverage reports. Finally, you will learn how to deploy your application on Heroku and AWS using various technologies, such as Docker, CloudFormation, and Elastic Beanstalk, and will also learn how to develop Jenkins pipelines to build, test, and deploy applications. What you will learn Develop a Flask extension using best practices Implement various authentication methods: LDAP, JWT, Database, OAuth, and OpenID Learn how to develop role-based access security and become an expert on Jinja2 templates Build tests for

your applications and APIsInstall and configure a distributed task queue using Celery and RabbitMQDevelop RESTful APIs and secure REST APIsDeploy highly available applications that scale on Heroku and AWS using Docker or VMsWho this book is for The ideal target audience for this book would be Python developers who want to use Flask and its advanced features to create Enterprise grade and lightweight applications. The book is for those who have some exposure of Flask and want to take it from introductory to master level.

The second edition of the best-selling Python book in the world (over 1 million copies sold!). A fast-paced, no-nonsense guide to programming in Python. Updated and thoroughly revised to reflect the latest in Python code and practices. Python Crash Course is the world's best-selling guide to the Python programming language. This fast-paced, thorough introduction to programming with Python will have you writing programs, solving problems, and making things that work in no time. In the first half of the book, you'll learn basic programming concepts, such as variables, lists, classes, and loops, and practice writing clean code with exercises for each topic. You'll also learn how to make your programs interactive and test your code safely before adding it to a project. In the second half, you'll put your new knowledge into practice with three substantial projects: a Space Invaders-inspired arcade game, a set of data visualizations with Python's handy libraries, and a simple web app you can deploy online. As you work through the book, you'll learn how to:

- Use powerful Python libraries and tools, including Pygame, Matplotlib, Plotly, and Django
- Make 2D games that respond to keypresses and mouse clicks, and that increase in difficulty
- Use data to generate interactive visualizations
- Create and customize web apps and deploy them safely online
- Deal with mistakes and errors so you can solve your own programming problems if you've been thinking about digging into programming. Python Crash Course will get you writing real programs fast. Why wait any longer? Start your engines and code!

Gray Hat Python

Python All-in-One For Dummies

Mastering Flask Web Development

Python Web Scraping

Web Scraping with Python

Data Science from Scratch

Flask Web Development

Build a solid foundation in coding by utilizing the language and its core characteristics **Key Features** **Leverage the features of Python programming through easy-to-follow examples** **Develop a strong set of programming skills that can be applied on all platforms** **Create GUIs and data science-based applications** **Book Description** **Learn Python Programming** creates a foundation for those who are interested in developing their skills in Python programming. The book starts with the fundamentals of programming with Python and ends by exploring different topics such as GUIs and real-world apps. You will begin by exploring the foundations of and fundamental topics on Python and learn to manipulate them. Then, you'll explore different programming paradigms that will allow you to find the best approach to a situation, and you'll also understand how to carry out performance optimization as well as effective debugging. As you make your way through the chapters, you'll control the flow of a program, and persist and utilize an interchange format to exchange data. You'll also walk through cryptographic services in Python and understand secure tokens. Throughout, the book covers various types of applications, and it concludes with building real-world applications based on all the concepts that you learned. By the end of the book, you'll have a proper understanding of the Python language and a solid grasp on how to work with data. You'll know how to quickly build a website and harness the power of Python's renowned data science libraries. What you will learn Get Python up and running on Windows, Mac, and Linux Grasp fundamental concepts of coding using data structures and control flow Write elegant, reusable, and efficient code in any situation Understand when to use the functional or object-oriented programming (OOP) approach Walk through the basics of security and concurrent/asynchronous programming Create bulletproof, reliable software by writing tests Explore examples of GUIs, scripting, and data science Who this book is for Learn Python Programming is for individuals with relatively little experience in coding or Python. It's also ideal for aspiring programmers who need to write scripts or programs to accomplish tasks. The book takes you all the way to creating a full-fledged application.

Easy to understand and fun to read, this updated edition of Introducing Python is ideal for beginning programmers as well as those new to the language. Author Bill Lubanovic takes you from the basics to more involved and varied topics, mixing tutorials with cookbook-style code recipes to explain concepts in Python 3. **End-of-chapter exercises help you practice what you've learned. You'll gain a strong foundation in the language, including best practices for testing, debugging, code reuse, and other development tips.** This book also shows you how to use Python for applications in business, science, and the arts, using various Python tools and open source packages.

Take full creative control of your web applications with Flask, the Python-based microframework. With the second edition of this hands-on book, you'll learn the framework from the ground up by developing, step-by-step, a real-world project created by author Miguel Grinberg. This refreshed edition accounts for important technology changes that have occurred in the past three years. You'll learn the framework's core functionality, as well as how to extend applications with advanced web techniques such as database migration and web service communication. The first part of each chapter provides you with references to take you. These 7 straightforward and friendly mini-books assume the reader is a beginning programmer, and cover everything from the basic elements of Python code to introductions to the specific applications where you'll use it. Intended as a hands-on reference, the focus is on practice over theory, providing you with examples to follow as well as code for you to copy and start modifying in the "real world"—helping you get up and running in your area of interest almost right away. This means you'll be finishing off your first app or building and remote-controlling your own robot much faster than you can believe. Get a thorough grounding in the language basics Learn how syntax is applied in high-profile industries Apply Python to projects in enterprise Find out how Python can get you into hot careers in AI, big data, and more Whether you're a newbie coder or just want to add Python to your magic box of tricks, this is the perfect, practical introduction—and one you'll return to as you grow your career.

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.pythontutorial.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 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. 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 a 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

Best Practices for Development

The no-nonsense, beginner's guide to programming, data science, and web development with Python 3.7, 2nd Edition

Python Web Scraping Cookbook

Learn Python Programming

Practical Web Scraping for Data Science

Exploring Data in Python 3

Over 90 proven recipes to get you scraping with Python, microservices, Docker, and AWS

Your Python code may run correctly, but you need it to run faster. Updated for Python 3, this expanded edition shows you how to locate performance bottlenecks and significantly speed up your code in high-data-volume programs. By exploring the fundamental theory behind design choices, High Performance Python helps you gain a deeper understanding of Python implementation. How do you take advantage of multicore architectures or clusters? Or build a system that scales up and down without losing reliability? Experienced Python programmers will learn concrete solutions to many issues, along with war stories from companies that use high-performance Python for social media analytics, productionized machine learning, and more. Get a better grasp of NumPy, Cython, and profilers Learn how Python abstracts the underlying computer architecture Use profiling to find bottlenecks in CPU time and memory usage Write efficient programs by choosing appropriate data structures Speed up matrix and vector computations Use tools to compile Python down to machine code Manage mult I/O and computational operations concurrently Convert multiprocessing code to run on local or remote clusters Deploy code faster using tools like Docker

Web Scraping techniques are getting more popular, since data is as valuable as oil in 21st century. Through this book get some key knowledge about using XPath, regex, web scraping libraries for R like rvest and RSelenium technologies. Key FeaturesTechniques, tools and frameworks for web scraping with RScrape data effortlessly from a variety of websites Learn how to selectively choose the data to scrape, and build your datasetBook Description Web scraping is a technique to extract data from websites. It simulates the behavior of a website user to turn the website itself into a web service to retrieve or introduce new data. This book gives you all you need to get started with scraping web pages using R programming. You will learn about the rules of RegEx and Xpath, key components for scraping website data. We will show you web scraping techniques, methodologies, and frameworks. With this book's guidance, you will become comfortable with the tools to write and test RegEx and XPath rules. We will focus on examples of dynamic websites for scraping data and how to implement them using the techniques learned. You will learn how to collect URLs and then create XPath rules for your first web scraping script using rvest library. From the data you collect, you will be able to calculate the statistics and create R plots to visualize them. Finally, you will discover how to use Selenium drivers with R for more sophisticated scraping. You will create AWS instance use R to connect a PostgreSQL database hosted on AWS. By the end of the book, you will be sufficiently confident to create end-to-end web scraping systems using R. What you will learnWrite and create regex rulesWrite XPath rules to query your dataLearn how web scraping methods workUse R to crawl web pagesStore data retrieved from the webLearn the uses of Relenium to scrape dataWho this book is for This book is for R programmers who want to get started quickly with web scraping, as well as data analysts who want to learn scraping using R. Basic knowledge of R is all you need to get started with this book.

Immerse yourself in learning Python and introductory data analytics with this book's project-based approach. Through the structure of a ten-week coding bootcamp course, you'll learn key concepts and gain hands-on experience through weekly projects. Each chapter in this book is presented as a full week of topics, with Monday through Thursday covering specific concepts, leading up to Friday, when you are challenged to create a project using the skills learned throughout the week. Topics include Python basics and essential intermediate concepts such as list comprehension, generators and iterators, understanding algorithmic complexity, and data analysis with pandas. From beginning to end, this book builds up your abilities through exercises and challenges, culminating in your solid understanding of Python. Challenge yourself with the intensity of a coding bootcamp experience or learn at your own pace. With this hands-on learning approach, you will gain the skills you need to jumpstart a new career in programming or further your current one as a software developer. What You Will Learn Understand beginning and more advanced concepts of the Python languageBe introduced to data analysis using pandas, the Python Data Analysis libraryWalk through the process of interviewing and answering technical questionsCreate real-world applications with the Python languageLearn how to use Anaconda, Jupyter Notebooks, and the Python Shell Who This Book Is For Those trying to jumpstart a new career into programming, and those already in the software development industry and would like to learn Python programming.

Python's simplicity lets you become productive quickly, but this often means you aren't using everything it has to offer. With this hands-on guide, you'll learn how to write effective, idiomatic Python code by leveraging its best—and possibly most neglected—features. Author Luciano Ramalho takes you through Python's core language features and libraries, and shows you how to make your code shorter, faster, and more readable at the same time. Many experienced programmers try to bend Python to fit patterns they learned from other languages, and never discover Python features outside of their experience. With this book, those Python programmers will thoroughly learn how to become proficient in Python 3. This book covers Python data model, understand how special methods are the key to the consistent behavior of objects Data structures: take full advantage of built-in types, and understand the text vs bytes duality in the Unicode age Functions as objects: view Python functions as first-class objects, and understand how this affects popular design patterns Object-oriented idioms: classes by learning about references, mutability, interfaces, operator overloading, and multiple inheritance Control flow: leverage context managers, generators, coroutines, and concurrency with the concurrent futures and asyncio packages Metaprogramming: understand how properties, attribute descriptors, class decorators, and metaclasses work

Practical Programming for Total Beginners

Python Automation Cookbook

Python Web Scraping, Second Edition

Collecting Data from the Modern Web

Data Science For Dummies

A complete guide to pandas, from installation to advanced data analysis techniques, 2nd Edition