

## Package Apt R

\* An indispensable resource for Fedora users who must now work without customer support from Red Hat, Inc., covering critical troubleshooting techniques for networks, internal servers, and external servers
\* Chris Negus is a well-known Linux authority and also the author of the top-selling Red Hat Linux Bible (0-7645-4333-4); Thomas Weeks is a trainer and administrator who manages hundreds of Red Hat Linux systems
\* Covers all of the most common Fedora problem areas: firewalls, DNS servers, print servers, Samba, NFS, Web servers, FTP servers, e-mail servers, modems, adding hardware, and hardware certification
\* Features easy-to-use flowcharts that guide administrators step by step through common Fedora troubleshooting scenarios
\* A companion Web site offers troubleshooting updates to keep pace with the frequent Fedora Core releases as well as a forum for exchanging troubleshooting tips

Games provide mathematical models for interaction. Numerous tasks in computer science can be formulated in game-theoretic terms. This fresh and intuitive way of thinking through complex issues reveals underlying algorithmic questions and clarifies the relationships between different domains. This collection of lectures, by specialists in the field, provides an excellent introduction to various aspects of game theory relevant for applications in computer science that concern program design, synthesis, verification, testing and design of multi-agent or distributed systems. Originally devised for a Spring School organised by the GAMES Networking Programme in 2009, these lectures have since been revised and expanded, and range from tutorials concerning fundamental notions and methods to more advanced presentations of current research topics. This volume is a valuable guide to current research on game-based methods in computer science for undergraduate and graduate students. It will also interest researchers working in mathematical logic, computer science and game theory.

As more and more organizations migrate their applications to the cloud, cloud native computing has become the dominant way to approach software development and execution. Protecting modern, cloud native applications from threats requires the ability to defend them at runtime, when they’re most vulnerable to attacks. This practical guide introduces you to Falco, the open source standard for continuous risk and threat detection across Kubernetes, containers, and the cloud. Falco creator Loris Degioanni and core maintainer Leonardo Grasso bring you up to speed on threat detection and show you how to get Falco up and running, plus advanced topics such as deploying Falco in production and writing your own security rules. You'll learn how to: Leverage runtime security in cloud native environments Detect configuration changes and unexpected behavior in the cloud Protect containers, Kubernetes, and cloud applications using Falco Run, deploy, and customize Falco Deploy, configure, and maintain Falco in a production environment Improve your compliance

R Markdown: The Definitive Guide is the first official book authored by the core R Markdown developers that provides a comprehensive and accurate reference to the R Markdown ecosystem. With R Markdown, you can easily create reproducible data analysis reports, presentations, dashboards, interactive applications, books, dissertations, websites, and journal articles, while enjoying the simplicity of Markdown and the great power of R and other languages. In this book, you will learn Basics: Syntax of Markdown and R code chunks, how to generate figures and tables, and how to use other computing languages Built-in output formats of R Markdown: PDF/HTML/Word/RTF/Markdown documents and ioslides/Slidy/Beamer/PowerPoint presentations Extensions and applications: Dashboards, Tufte handouts, xaringan/reveal.js presentations, websites, books, journal articles, and interactive tutorials Advanced topics: Parameterized reports, HTML widgets, document templates, custom output formats, and Shiny documents. Yihui Xie is a software engineer at RStudio. He has authored and co-authored several R packages, including knitr, rmarkdown, bookdown, blogdown, shiny, xaringan, and animation. He has published three other books, Dynamic Documents with R and knitr, bookdown: Authoring Books and Technical Documents with R Markdown, and blogdown: Creating Websites with R Markdown. J.J. Allaire is the founder of RStudio and the creator of the RStudio IDE. He is an author of several packages in the R Markdown ecosystem including rmarkdown, flexdashboard, learnr, and radix. Garrett Grolmund is the co-author of R for Data Science and author of Hands-On Programming with R. He wrote the lubridate R package and works for RStudio as an advocate who trains engineers to do data science with R and the Tidyverse.

Geocomputation with R

An R Companion to Applied Regression

Advanced R

Machine Learning with Amazon SageMaker Cookbook

Administration and Reference

Predictive Soil Mapping with R

This book is designed as an Ubuntu 20.04 LTS Server administration and reference source, covering the Ubuntu servers and their support applications. Server tools are covered as well as the underlying configuration files and system implementations. The emphasis is on what administrators will need to know to perform key server support and management tasks. Coverage of the systemd service management system is integrated into the book. Topics covered include software management, systemd service management, systemd-networkd and Netplan network configuration, AppArmor security, OpenSSH, the Chrony time server, and Ubuntu cloud services. Key servers are examined, including Web, FTP, CUPS printing, NFS, and Samba Windows shares. Network support servers and applications covered include the Squid proxy server, the Domain Name System (BIND) server, DHCP, distributed network file systems, IPTables firewalls, and cloud computing.

This document is designed to be a resource for those Linux users wishing to seek clarification on Linux/UNIX/POSIX related terms and jargon. At approximately 24000 definitions and two thousand pages it is one of the largest Linux related dictionaries currently available. Due to the rapid rate at which new terms are being created it has been decided that this will be an active project. We welcome input into the content of this document. At this moment in time half yearly updates are being envisaged. Please note that if you wish to find a 'Computer Dictionary' then see the 'Computer Dictionary Project' at http://computerdictionary.tsf.org.za/ Searchable databases exist at locations such as: http://www.swpearl.com/eng/scripts/dictionary/ (SWP) Sun Wah-PearL Linux Training and Development Centre is a centre of the Hong Kong Polytechnic University, established in 2000. Presently SWP is delivering professional grade Linux and related Open Source Software (OSS) technology training and consultant service in Hong Kong. SWP has an ambitious aim to promote the use of Linux and related Open Source Software (OSS) and Standards. The vendor independent positioning of SWP has been very well perceived by the market. Throughout the last couple of years, SWP becomes the Top Leading OSS training and service provider in Hong Kong. http://www.geona.com/dictionary?b= Geona, operated by Gold Vision Communications, is a new powerful search engine and internet directory, delivering quick and relevant results on almost any topic or subject you can imagine. The term "Geona" is an Italian and Hebrew name, meaning wisdom, exaltation, pride or majesty. We use our own database of spidered web sites and the Open Directory database, the same database which powers the core directory services for the Web's largest and most popular search engines and portals. Geona is spidering all domains listed in the non-adult part of the Open Directory and millions of additional sites of general interest to maintain a fullest index of highly relevant web sites. http://www.linuxdig.com/documents/dictionary.php LINUXDIG.COM, "Yours News and Resource Site", LinuxDig.com was started in May 2001 as a hobby site with the original intention of getting the RFC's online and becoming an Open Source software link/download site. But since that time the site has evolved to become a RFC distribution site, linux news site and a locally written technology news site (with bad grammer :) with focus on Linux while also containing articles about anything and everything we find interesting in the computer world. LinuxDig.Com contains about 20,000 documents and this number is growing everyday! http://linux.about.com/library/glossary/bglossary.htm Each month more than 20 million people visit About.com. Whether it be home repair and decorating ideas, recipes, movie trailers, or car buying tips, our Guides offer practical advice and solutions for every day life. Wherever you land on the new About.com, you'll find other content that is relevant to your interests. If you're looking for "How To" advice on planning to re-finish your deck, we'll also show you the tools you need to get the job done. If you've been to About before, we'll show you the latest updates, so you don't see the same thing twice. No matter where you are on About.com, or how you got here, you'll always find content that is relevant to your needs. Should you wish to possess your own localised searchable version please make use of the available "dict", http://www.dict.org/ version at the Linux Documentation Project home page, http://www.tldp.org/ The author has decided to leave it up to readers to determine how to install and run it on their specific systems. An alternative form of the dictionary is available at: http://elibrary.fultus.com/covers/technical/linux/guides/Linux-Dictionary/cover.html Fultus Corporation helps writers and companies to publish, promote, market, and sell books and eBooks. Fultus combines traditional self-publishing practices with modern technology to produce paperback and hardcover print-on-demand (POD) books and electronic books (eBooks). Fultus publishes works (fiction, non-fiction, science fiction, mystery, ...) by both published and unpublished authors. We enable you to self-publish easily and cost-effectively, creating your book as a print-ready paperback or hardcover POD book or as an electronic book (eBook) in multiple eBook's formats. You retain all rights to your work. We provide distribution to bookstores worldwide. And all at a fraction of the cost of traditional publishing. We also offer corporate publishing solutions that enable businesses to produce and deliver manuals and documentation more efficiently and economically. Our use of electronic delivery and print-on-demand technologies reduces printed inventory and saves time. Please inform the author as to whether you would like to create a database or an alternative form of the dictionary so that he can include you in this list. Also note that the author considers breaches of copyright to be extremely serious. He will pursue all claims to the fullest extent of the law.

Thorough LPIC-1 exam prep, with complete coverage and bonusstudy tools LPIC-1Study Guide is your comprehensive source for thepopular Linux Professional Institute Certification Level 1 exam,fully updated to reflect the changes to the latest version of theexam. With 100% coverage of objectives for both LPI 101 and LPI102, this book provides clear and concise information on all Linuxadministration topics and practical examples drawn from real-worldexperience. Authoritative coverage of key exam topics includes GNUand UNIX commands, devices, file systems, file system hierarchy,user interfaces, and much more, providing complete exam prep forthe LPIC-1 candidate. Get access to invaluable study tools,including bonus practice exams, electronic flashcards, and aseachable PDF of key terms featured on the exam. Linux is viewed by many companies and organizations as anexcellent, low-cost, secure alternative to expensive operatingsystems, such as Microsoft Windows. The LPIC-1 tests a candidate'sunderstanding and familiarity with the Linux Kernel. This bookprovides comprehensive preparation and review, helping readers facethe exam with confidence. Review the system architecture, Linux installation, and packagemanagement Understand shells, scripting, and data management morecompletely Practice administrative tasks and essential systemservices Brush up on networking fundamentals and security issues As the Linux server market share continue to grow, so too doesthe demand for qualified and certified Linux administrators.Certification holders must recertify every five years, but LPRecommends recertifying every two years to stay fully up to datewith new technologies and best practices. As exam day approaches,LPIC-1Study Guide is the one source you will want by yourside.

Europe is getting closer. So are European social sciences. However, this is easier done in theory development and central research questions. When it comes to data the mutual understanding is far from perfect, due to a lack of knowledge about the data bases of the respective countries and the EU in general. This is particularly true when it comes to the regional level. This volume will help to improve the insight into the rich stock of European datasets which cover any kind of regional information. Many institutions ranging from statistical offices to more academic research centres and commercial enterprises report their offerings with special emphasis on the regional level (e.g. European Community Household Panel, European Social Survey, Labour Force Survey). Central categories such as NUTS and LAU are explained and discussed. In addition, typical examples of socio-economic cross-border and multi-level studies highlight the power of a regionalized European perspective. Furthermore, information about special tools for such type of analysis is included in the volume.

LPIC-1: Linux Professional Institute Certification Study Guide

Engineering Production-Grade Shiny Apps

A Practical Guide to NCBI Databases and Sequence Alignments

CompTIA Linux+ Complete Study Guide Authorized Courseware

Thinking with Examples for Effective Learning

Lectures in Game Theory for Computer Scientists

**Over 100 recipes to get up and running with the modern Linux administration ecosystem**
**Key Features**
**Understand and implement the core system administration tasks in Linux**
**Discover tools and techniques to troubleshoot your Linux system**
**Maintain a healthy system with good security and backup practices**
**Book Description**
**Linux is one of the most widely used operating systems among system administrators, and even modern application and server development is heavily reliant on the Linux platform. The Linux Administration Cookbook is your go-to guide to get started on your Linux journey. It will help you understand what that strange little server is doing in the corner of your office, what the mysterious virtual machine languishing in Azure is crunching through, what that circuit-board-like thing is doing under your office TV, and why the LEDs on it are blinking rapidly. This book will get you started with administering Linux, giving you the knowledge and tools you need to troubleshoot day-to-day problems, ranging from a Raspberry Pi to a server in Azure, while giving you a good understanding of the fundamentals of how GNU/Linux works. Through the course of the book, you'll install and configure a system, while the author regales you with errors and anecdotes from his vast experience as a data center hardware engineer, systems administrator, and DevOps consultant. By the end of the book, you will have gained practical knowledge of Linux, which will serve as a bedrock for learning Linux administration and aid you in your Linux journey. What you will learn**
**Install and manage a Linux server, both locally and in the cloud**
**Understand how to perform administration across all Linux distros**
**Work through evolving concepts such as IaaS versus PaaS, containers, and automation**
**Explore security and configuration best practices**
**Troubleshoot your system if something goes wrong**
**Discover and mitigate hardware issues, such as faulty memory and failing drives**
**Who this book is for**
**If you are a system engineer or system administrator with basic experience of working with Linux, this book is for you.**

**Build, manage, and configure high-performing, reliable NoSQL database for your applications with Cassandra**
**Key Features**
**Write programs more efficiently using Cassandra's features with the help of examples**
**Configure Cassandra and fine-tune its parameters depending on your needs**
**Integrate Cassandra database with Apache Spark and build strong data analytics pipeline**
**Book Description**
**With ever-increasing rates of data creation, the demand for storing data fast and reliably becomes a need. Apache Cassandra is the perfect choice for building fault-tolerant and scalable databases. Mastering Apache Cassandra 3.x teaches you how to build and architect your clusters, configure and work with your nodes, and program in a high-throughput environment, helping you understand the power of Cassandra as per the new features. Once you've covered a brief recap of the basics, you'll move on to deploying and monitoring a production setup and optimizing and integrating it with other software. You'll work with the advanced features of CQL and the new storage engine in order to understand how they function on the server-side. You'll explore the integration and interaction of Cassandra components, followed by discovering features such as token allocation algorithm, CQL3, vnodes, lightweight transactions, and data modelling in detail. Last but not least you will get to grips with Apache Spark. By the end of this book, you'll be able to analyse big data, and build and manage high-performance databases for your application. What you will learn**
**Write programs more efficiently using Cassandra's features more efficiently**
**Exploit the given infrastructure, improve performance, and tweak the Java Virtual Machine (JVM)**
**Use CQL3 in your application in order to simplify working with Cassandra**
**Configure Cassandra and fine-tune its parameters depending on your needs**
**Set up a cluster and learn how to scale it**
**Monitor a Cassandra cluster in different ways**
**Use Apache Spark and other big data processing tools**
**Who this book is for**
**Mastering Apache Cassandra 3.x is for you if you are a big data administrator, database administrator, architect, or developer who wants to build a high-performing, scalable, and fault-tolerant database. Prior knowledge of core concepts of databases is required.**

**Over insightful 90 recipes to get lightning-fast analytics with Apache Spark**
**About This Book**
**Use Apache Spark for data processing with these hands-on recipes**
**Implement end-to-end, large-scale data analysis better than ever before**
**Work with powerful libraries such as MLlib, SciPy, NumPy, and Pandas to gain insights from your data**
**Who This Book Is For**
**This book is for novice and intermediate level data science professionals and data analysts who want to solve data science problems with a distributed computing framework. Basic experience with data science implementation tasks is expected. Data science professionals looking to skill up and gain an edge in the field will find this book helpful. What You Will Learn**
**Explore the topics of data mining, text mining, Natural Language Processing, information retrieval, and machine learning. Solve real-world analytical problems with large data sets. Address data science challenges with analytical tools on a distributed system like Spark (apt for iterative algorithms), which offers in-memory processing and more flexibility for data analysis at scale. Get hands-on experience with algorithms like Classification, regression, and recommendation on real datasets using Spark MLlib package. Learn about numerical and scientific computing using NumPy and SciPy on Spark. Use Predictive Model Markup Language (PMML) in Spark for statistical data mining models. In Detail**
**Spark has emerged as the most promising big data analytics engine for data science professionals. The true power and value of Apache Spark lies in its ability to execute data science tasks with speed and accuracy. Spark's selling point is that it combines ETL, batch analytics, real-time stream analysis, machine learning, graph processing, and visualizations. It lets you tackle the complexities that come with raw unstructured data sets with ease. This guide will get you comfortable and confident performing data science tasks with Spark. You will learn about implementations including distributed deep learning, numerical computing, and scalable machine learning. You will be shown effective solutions to problematic concepts in data science using Spark's data science libraries such as MLlib, Pandas, NumPy, SciPy, and more. These simple and efficient recipes will show you how to implement algorithms and optimize your work. Style and approach**
**This book contains a comprehensive range of recipes designed to help you learn the fundamentals and tackle the difficulties of data science. This book outlines practical steps to produce powerful insights into Big Data through a recipe-based approach.**

Geocomputation with RCRC Press

Stochastic Modelling for Systems Biology, Third Edition

Apache Spark for Data Science Cookbook

Europe and its Regions

Learn Linux Quickly

Linux Troubleshooting Bible

Bioinformatics

Geocomputation with R is for people who want to analyze, visualize and model geographic data with open source software. It is based on R, a statistical programming language that has powerful data processing, visualization, and geospatial capabilities. The book equips you with the knowledge and skills to tackle a wide range of issues manifested in geographic data, including those with scientific, societal, and environmental implications. This book will interest people from many backgrounds, especially Geographic Information Systems (GIS) users interested in applying their domain-specific knowledge in a powerful open source language for data science, and R users interested in extending their skills to handle spatial data. The book is divided into three parts: (I) Foundations, aimed at getting you up-to-speed with geographic data in R, (II) extensions, which covers advanced techniques, and (III) applications to real-world problems. The chapters cover progressively more advanced topics, with early chapters providing strong foundations on which the later chapters build. Part I describes the nature of spatial datasets in R and methods for manipulating them. It also covers geographic data import/export and transforming coordinate reference systems. Part II represents methods that build on these foundations. It covers advanced map making (including web mapping), "bridges" to GIS, sharing reproducible code, and how to do cross-validation in the presence of spatial autocorrelation. Part III applies the knowledge gained to tackle real-world problems, including representing and modeling transport systems, finding optimal locations for stores or services, and ecological modeling. Exercises at the end of each chapter give you the skills needed to tackle a range of geospatial problems. Solutions for each chapter and supplementary materials providing extended examples are available at https://geocompr.github.io/geocompkg/articles/. Dr. Robin Lovelace is a University Academic Fellow at the University of Leeds, where he has taught R for geographic research over many years, with a focus on transport systems. Dr. Jakub Nowosad is an Assistant Professor in the Department of Geoinformation at the Adam Mickiewicz University in Poznan, where his focus is on the analysis of large datasets to understand environmental processes. Dr. Jannes Muenchow is a Postdoctoral Researcher in the GIScience Department at the University of Jena, where he develops and teaches a range of geographic methods, with a focus on ecological modeling, statistical geocomputing, and predictive mapping. All three are active developers and work on a number of R packages, including stplanr, sabre, and RQGIS.

If you have always wanted to learn Linux but are still afraid to do so, this book is for you! A lot of people think of Linux as a sophisticated operating system that only hackers and geeks know how to use, and thus they abort their dream of learning Linux. Well, let me surprise you! Linux is simple and easy to learn, and this book is the ultimate proof! You may have stumbled across a variety of sources that all explain Linux in a complicated and dry manner. This book does exactly the opposite; it teaches you Linux in a delightful and friendly way so that you will never get bored, and you will always feel motivated to learn more. Learn Linux Quickly doesn't assume any prior Linux knowledge, which makes it a perfect fit for beginners. Nevertheless, intermediate and advanced Linux users will still find this book very useful as it goes through a wide range of topics. Learn Linux Quickly will teach you the following topics: · Installing Linux · Over 116 Linux Commands · User and Group Management · Linux Networking Fundamentals · Bash Scripting · Automate Boring Tasks with Cron Jobs · Create your Own Linux Commands · Linux Disk Partitioning and LVM · Finding Files on Linux · Understanding File Permissions · Linux Processes And much more! There is no time to waste here! Learn Linux Quickly and kick start your Linux career today!

Satellite Earth observation (EO) data have already exceeded the petabyte scale and are increasingly freely and openly available from different data providers. This poses a number of issues in terms of volume (e.g., data volumes have increased 10× in the last 5 years); velocity (e.g., Sentinel-2 is capturing a new image of any given place every 5 days); and variety (e.g., different types of sensors, spatial/spectral resolutions). Traditional approaches to the acquisition, management, distribution, and analysis of EO data have limitations (e.g., data size, heterogeneity, and complexity) that impede their true information potential to be realized. Addressing these big data challenges requires a change of paradigm and a move away from local processing and data distribution methods to lower the barriers caused by data size and related complications in data management. To tackle these issues, EO data cubes (EODC) are a new paradigm revolutionizing the way users can store, organize, manage, and analyze EO data. This Special Issue is consequently aiming to cover the most recent advances in EODC developments and implementations to broaden the use of EO data to larger communities of users, support decision-makers with timely and actionable information converted into meaningful geophysical variables, and ultimately unlock the information power of EO data.

Machine learning has finally come of age. With H2O software, you can perform machine learning and data analysis using a simple open source framework that’s easy to use, has a wide range of OS and language support, and scales for big data. This hands-on guide teaches you how to use H2O with only minimal math and theory behind the learning algorithms. If you’re familiar with R or Python, know a bit of statistics, and have some experience manipulating data, author Darren Cook will take you through H2O basics and help you conduct machine-learning experiments on different sample data sets. You’ll explore several modern machine-learning techniques such as deep learning, random forests, unsupervised learning, and ensemble learning. Learn how to import, manipulate, and export data with H2O Explore key machine-learning concepts, such as cross-validation and validation data sets Work with three diverse data sets, including a regression, a multinomial classification, and a binomial classification Use H2O to analyze each sample data set with four supervised machine-learning algorithms Understand how cluster analysis and other unsupervised machine-learning algorithms work

Interpretable Machine Learning

LPIC-1 Linux Professional Institute Certification Study Guide  
 Insightful recipes to work with system administration tasks on Linux  
 Exam 101-400 and Exam 102-400  
 Machine Learning Models and Algorithms for Big Data Classification  
 (Exams 101 and 102)

This book presents machine learning models and algorithms to address big data classification problems. Existing machine learning techniques like the decision tree (a hierarchical approach), random forest (an ensemble hierarchical approach), and deep learning (a layered approach) are highly suitable for the system that can handle such problems. This book helps readers, especially students and newcomers to the field of big data and machine learning, to gain a quick understanding of the techniques and technologies; therefore, the theory, examples, and programs (Matlab and R) presented in this book have been simplified, hardcoded, repeated, or spaced for improvements. They provide vehicles to test and understand the complicated concepts of various topics in the field. It is expected that the readers adopt these programs to experiment with the examples, and then modify or write their own programs toward advancing their knowledge for solving more complex and challenging problems. The presentation format of this book focuses on simplicity, readability, and dependability so that both undergraduate and graduate students as well as new researchers, developers, and practitioners in this field can easily trust and grasp the concepts, and learn them effectively. It has been written to reduce the mathematical complexity and help the vast majority of readers to understand the topics and get interested in the field. This book consists of four parts, with the total of 14 chapters. The first part mainly focuses on the topics that are needed to help analyze and understand data and big data. The second part covers the topics that can explain the systems required for processing big data. The third part presents the topics required to understand and select machine learning techniques to classify big data. Finally, the fourth part concentrates on the topics that explain the scaling-up machine learning, an important solution for modern big data problems.

There are many excellent R resources for visualization, data science, and package development. Hundreds of scattered vignettes, web pages, and forums explain how to use R in particular domains. But little has been written on how to simply make R work effectively—until now. This hands-on book teaches novices and experienced R users how to write efficient R code. Drawing on years of experience teaching R courses, authors Colin Gillespie and Robin Lovelace provide practical advice on a range of topics—from optimizing the set-up of RStudio to leveraging C++—that make this book a useful addition to any R user’s bookshelf. Academics, business users, and programmers from a wide range of backgrounds stand to benefit from the guidance in Efficient R Programming. Get advice for setting up an R programming environment Explore general programming concepts and R coding techniques Understand the ingredients of an efficient R workflow Learn how to efficiently read and write data in R Dive into data carpentry—the vital skill for cleaning raw data Optimize your code with profiling, standard tricks, and other methods Determine your hardware capabilities for handling R computation Maximize the benefits of collaborative R programming Accelerate your transition from R hacker to R programmer

Advance your skills in efficient data analysis and data processing using the powerful tools of Scala, Spark, and Hadoop About This Book This is a primer on functional-programming-style techniques to help you efficiently process and analyze all of your data Get acquainted with the best and newest tools available such as Scala, Spark, Parquet and MLlib for machine learning Learn the best practices to incorporate new Big Data machine learning in your data-driven enterprise to gain future scalability and maintainability Who This Book Is For Mastering Scala Machine Learning is intended for enthusiasts who want to plunge into the new pool of emerging techniques for machine learning. Some familiarity with standard statistical techniques is required. What You Will Learn Sharpen your functional programming skills in Scala using REPL Apply standard and advanced machine learning techniques using Scala Get acquainted with Big Data technologies and grasp why we need a functional approach to Big Data Discover new data structures, algorithms, approaches, and habits that will allow you to work effectively with large amounts of data Understand the principles of supervised and unsupervised learning in machine learning Work with unstructured data and serialize it using Kryo, Protobuf, Avro, and AvroParquet Construct reliable and robust data pipelines and manage data in a data-driven enterprise Implement scalable model monitoring and alerts with Scala In Detail Since the advent of object-oriented programming, new technologies related to Big Data are constantly popping up on the market.

One such technology is Scala, which is considered to be a successor to Java in the area of Big Data by many, like Java was to C/C++ in the area of distributed programing. This book aims to take your knowledge to next level and help you impart that knowledge to build advanced applications such as social media mining, intelligent news portals, and more. After a quick refresher on functional programming concepts using REPL, you will see some practical examples of setting up the development environment and tinkering with data. We will then explore working with Spark and MLlib using k-means and decision trees. Most of the data that we produce today is unstructured and raw, and you will learn to tackle this type of data with advanced topics such as regression, classification, integration, and working with graph algorithms. Finally, you will discover at how to use Scala to perform complex concept analysis, to monitor model performance, and to build a model repository. By the end of this book, you will have gained expertise in performing Scala machine learning and will be able to build complex machine learning projects using Scala. Style and approach This hands-on guide dives straight into implementing Scala for machine learning without delving much into mathematical proofs or validations. There are ample code examples and tricks that will help you sail through using the standard techniques and libraries. This book provides practical examples from the field on how to correctly tackle data analysis problems, particularly for modern Big Data datasets.

From the Reviews "[This book] contains an excellent blend of both Shiny-specific topics ... and practical advice from software development that fits in nicely with Shiny apps. You will find many nuggets of wisdom sprinkled throughout these chapters...." Eric Nantz, Host of the R-Podcast and the Shiny Developer Series (from the Foreword) "[This] book is a gradual and pleasant invitation to the production-ready shiny apps world. It ...exposes a comprehensive and robust workflow powered by the {golem} package. [It] fills the not yet covered gap between shiny app development and deployment in such a thrilling way that it may be read in one sitting.... In the industry world, where processes robustness is a key toward productivity, this book will indubitably have a tremendous impact." David Granjon, Sr. Expert Data Science, Novartis Presented in full color, Engineering Production-Grade Shiny Apps helps people build production-grade shiny applications, by providing advice, tools, and a methodology to work on web applications with R. This book starts with an overview of the challenges which arise from any big web application project: organizing work, thinking about the user interface, the challenges of teamwork and the production environment. Then, it moves to a step-by-step methodology that goes from the idea to the end application. Each part of this process will cover in detail a series of tools and methods to use while building production-ready shiny applications. Finally, the book will end with a series of approaches and advice about optimizations for production. Features Focused on practical matters: This book does not cover Shiny concepts, but practical tools and methodologies to use for production. Based on experience: This book is a formalization of several years of experience building Shiny applications. Original content: This book presents new methodologies and tooling, not just a review of what already exists. Engineering Production-Grade Shiny Apps covers medium to advanced content about Shiny, so it will help people that are already familiar with building apps with Shiny, and who want to go one step further.

Engineering  
 Efficient R Programming  
 Ubuntu 20.04 LTS Server  
 Hacking with Kali  
 How to Build and Configure a Pbx With Open Source Software Featuring Relas 1.4

Bioinformatics derives knowledge from computer analysis of biological data. In particular, genomic and transcriptomic datasets are processed, analysed and, whenever possible, associated with experimental results from various sources, to draw structural, organizational, and functional information relevant to biology. Research in bioinformatics development for storage, retrieval, and analysis of the data. Bioinformatics in Aquaculture provides the most up to date reviews of next generation sequencing technologies, their applications in aquaculture, and principles and methodologies for the analysis of genomic and transcriptomic large datasets using bioinformatic methods, algorithmic unique in providing guidance for the best software packages suitable for various analysis, providing detailed examples of using bioinformatic software and command lines in the context of real world experiments. This book is a vital tool for all those working in genomics, molecular biology, biochemistry and genetics related to aquaculture, and sciences.

An Essential Reference for Intermediate and Advanced R Programmers Advanced R presents useful tools and techniques for attacking many types of R programming problems, helping you avoid mistakes and dead ends. With more than ten years of experience programming in R, the author illustrates the elegance, beauty, and flexibility at the develops the necessary skills to produce quality code that can be used in a variety of circumstances. You will learn: The fundamentals of R, including standard data types and functions Functional programming as a useful framework for solving wide classes of problems The positives and negatives of metaprogramming How to write fast, modular code not only helps current R users become R programmers but also shows existing programmers what’s special about R. Intermediate R programmers can dive deeper into R and learn new strategies for solving diverse problems while programmers from other languages can learn the details of R and understand why R works the way it does. CompTIA Authorized Linux+ prep CompTIA Linux+ Study Guide is your comprehensive study guide for the Linux+ Powered by LPI certification exams. With complete coverage of 100% of the objectives on both exam LX0-103 and exam LX0-104, this study guide provides clear, concise information on all aspects of Linux administration, with a focus on the of the exam. You’ll gain the insight of examples drawn from real-world scenarios, with detailed guidance and authoritative coverage of key topics, including GNU and Unix commands, system operation, system administration, system services, security, and more, from a practical perspective that easily translates into on-the-job know-how. You’ll also find study tools, including bonus practice exams, electronic flashcards, and a searchable glossary of key terms that are important to know for exam day. Linux is viewed by many companies and organizations as an excellent, low-cost, secure alternative to expensive operating systems such as Microsoft Windows. The CompTIA Linux+ Powered by LPI certification candidate’s understanding and familiarity with the Linux Kernel. Review the basic system architecture, installation, and management Understand commands, devices, and file systems Utilize shells, scripting, and data management techniques Navigate user interfaces, desktops, and essential system services As the Linux server market share continues to grow, it does the demand for qualified and certified Linux administrators. Certification holders must recertify every five years, but LPI recommends recertifying every two years to stay fully up to date with new technologies and best practices. CompTIA Linux+ Study Guide gives you the advantage of exam day confidence.

Since the first edition of Stochastic Modelling for Systems Biology, there have been many interesting developments in the use of "likelihood-free" methods of Bayesian inference for complex stochastic models. Having been thoroughly updated to reflect this, this third edition covers everything necessary for a good appreciation of stochastic modelling of biological networks in the systems biology context. New methods and applications are included in the book, and the use of R for practical illustration of the algorithms has been greatly extended. There is a brand new chapter on spatially extended systems, and the statistical inference chapter has also been extended with new methods, including approximate Bayesian computation (ABC). Stochastic Modelling for Systems Biology, Third Edition is now supplemented by an additional software library, written in Scala, described in a new appendix to the book. New in the Third Edition New chapter on spatially extended systems, covering the spatial Gillespie algorithm for reaction diffusion master equation models in 1- and 2-dimensions approximations based on the spatial chemical Langevin equation Significantly expanded chapter on inference for stochastic kinetic models from data, covering ABC, including ABC-SMC Updated R package, including code relating to all of the new material New R package for parsing SBML models into simulatable stochastic Petri net models New R package for parsing SBML models into simulatable stochastic Petri net models New R package for parsing SBML models into simulatable stochastic Petri net models New R library, written in Scala, replicating most of the functionality of the R packages in a fast, compiled, strongly typed, functional language Keeping with the spirit of earlier editions, all of the new theory is presented in a very informal and intuitive manner, keeping the text as accessible as possible to the widest possible readership. An effective and accessible introduction to stochastic modelling in computational systems biology, this new edition adds additional detail and computational methods that will provide a stronger foundation for the development of more advanced courses in stochastic biological modelling.

Practical Cloud Native Security with Falco  
 Linux Administration Cookbook  
 Learning Shiny  
 80 proven recipes for data scientists and developers to perform machine learning experiments and deployments  
 A Practical Guide to Smarter Programming  
 Mastering Scala Machine Learning

Make the most of R’s dynamic capabilities and implement web applications with Shiny About This Book Present interactive data visualizations in R within the Shiny framework Construct web dashboards in a simple, intuitive, but fully flexible environment Apply your skills to create a real-world web application with this step-by-step guide Who This Book Is For If you are a data scientist who needs a platform to show your results to a broader audience in an attractive and visual way, or a web developer with no prior experience in R or Shiny, this is the book for you. What You Will Learn Comprehend many useful functions, such as lapply and apply, to process data in R Write and structure different files to create a basic dashboard Develop graphics in R using popular graphical libraries such as ggplot2 and GoogleVis Mount a dashboard on a Linux Server Integrate Shiny with non-R-native visualization, such as D3.js Design and build a web application In Detail R is nowadays one of the most used tools in data science. However, along with Shiny, it is also gaining territory in the web application world, due to its simplicity and flexibility. Shiny is a framework that enables the creation of interactive visualizations written entirely in R and can be displayed in almost any ordinary web browser. It is a package from RStudio, which is an IDE for R. From the fundamentals of R to the administration of multi-concurrent, fully customized web applications, this book explains how to achieve your desired web application in an easy and gradual way. You will start by learning about the fundamentals of R, and will move on to looking at simple and practical examples. These examples will enable you to grasp many useful tools that will assist you in solving the usual problems that can be faced when developing data visualizations. You will then walk through the integration of Shiny with R in general and view the different visualization possibilities out there. Finally, you will put your skills to the test and create your first web application! Style and approach This is a comprehensive, step-by-step guide that will allow you to learn and make full use of R and Shiny’s capabilities in a gradual way, together with clear, applied examples.

Bioinformatics: A Practical Guide to NCBI Databases and Sequence Alignments provides the basics of bioinformatics and in-depth coverage of NCBI databases, sequence alignment, and NCBI Sequence Local Alignment Search Tool (BLAST). As bioinformatics has become essential for life sciences, the book has been written specifically to address the need of a large audience including undergraduates, graduates, researchers, healthcare professionals, and bioinformatics professors who need to use the NCBI databases, retrieve data from them, and use BLAST to find evolutionarily related sequences, sequence annotation, construction of phylogenetic tree, and the conservative domain of a protein, to name just a few. Technical details of alignment algorithms are explained with a minimum use of mathematical formulas and with graphical illustrations. Key Features Provides readers with the most-used bioinformatics knowledge of bioinformatics databases and alignments including both theory and application via illustrations and worked examples. Discusses the use of Windows Command Prompt, Linux shell, R, and Python for both Entrez databases and BLAST. The companion website contains tutorials, R and Python codes, instructor materials including slides, exercises, and problems for students. This is the ideal textbook for bioinformatics courses taken by students of life sciences and for researchers wishing to develop their knowledge of bioinformatics to facilitate their own research.

An R Companion to Applied Regression is a broad introduction to the R statistical computing environment in the context of applied regression analysis. John Fox and Sanford Weisberg provide a step-by-step guide to using the free statistical software R, an emphasis on integrating statistical computing in R with the practice of data analysis, coverage of generalized linear models, and substantial web-based support materials. The Third Edition has been reorganized and includes a new chapter on mixed-effects models, new and updated data sets, and a de-emphasis on statistical programming, while retaining a general introduction to basic R programming. The authors have substantially updated both the car and effects packages for R for this edition, introducing additional capabilities and making the software more consistent and easier to use. They also advocate an everyday data-analysis workflow that encourages reproducible research. To this end, they provide coverage of RStudio, an interactive development environment for R that allows readers to organize and document their work in a simple and intuitive fashion, and then easily share their results with others. Also included is coverage of R Markdown, showing how to create documents that mix R commands with explanatory text.

Ubuntu Server is a complete, free server operating system that just works, with the extra Ubuntu polish, innovation, and simplicity that administrators love. Now, there’s a definitive, authoritative guide to getting up and running quickly with the newest, most powerful versions of Ubuntu Server. Written by leading members of the Ubuntu community, The Official Ubuntu Server Book, Third Edition, covers all you need to know to make the most of Ubuntu Server, whether you’re a beginner or a battle-hardened senior systems administrator. The authors cover Ubuntu Server from start to finish: installation, basic administration and monitoring, security, backup, troubleshooting, system rescue, and much more. They walk through deploying each of the most common server applications, from file and print services to state-of-the-art, cost-saving virtualization and cloud computing. In addition, you’ll learn how to Make the most of Ubuntu Server’s latest, most powerful technologies Discover easy, fast ways to perform key administration tasks Automate Ubuntu installs, no matter how many servers you’re installing Quickly set up low-cost Web servers and e-mail Protect your server with Ubuntu’s built-in and optional security tools Minimize downtime with fault tolerance and clustering Master proven, step-by-step server and network troubleshooting techniques Walk through rescuing an Ubuntu server that won’t boot Deploy your own Ubuntu servers in the cloud

AUUGN  
 Principles and Methods  
 Mastering Apache Cassandra 3.x  
 The Official Ubuntu Server Book  
 Linux Dictionary  
 The Definitive Guide

**Predictive Soil Mapping (PSM) is based on applying statistical and/or machine learning techniques to fit models for the purpose of producing spatial and/or spatiotemporal predictions of soil variables i.e. maps of soil properties and classes at different resolutions. It is a multidisciplinary field combining statistics, data science, soil science, physical geography, remote sensing, geoinformation science and a number of other sciences. Predictive Soil Mapping with R is about understanding the main concepts behind soil mapping, mastering R packages that can be used to produce high quality soil maps, and about optimizing all processes involved so that also the production costs can be reduced. The online version of the book is available at: <https://envirometrix.github.io/PredictiveSoilMapping/> Pull requests and general comments are welcome. These materials are based on technical tutorials initially developed by the ISRIC’s Global Soil Information Facilities (GSIF) development team over the period 2014?2017**  
**Hacking with Kali introduces you the most current distribution of the de facto standard tool for Linux pen testing. Starting with use of the Kali live CD and progressing through installation on hard drives, thumb drives and SD cards, author James Broad walks you through creating a custom version of the Kali live distribution. You’ll learn how to configure networking components, storage devices and system services such as DHCP and web services. Once you’re familiar with the basic components of the software, you’ll learn how to use Kali through the phases of the penetration testing lifecycle; one major tool from each phase is explained. The book culminates with a chapter on reporting that will provide examples of documents used prior to, during and after the pen test. This guide will benefit information security professionals of all levels, hackers, systems administrators, network administrators, and beginning and intermediate professional pen testers, as well as students majoring in information security. Provides detailed explanations of the complete penetration testing lifecycle Complete linkage of the Kali information, resources and distribution downloads Hands-on exercises reinforce topics**  
**Learn how to construct machine learning and data analysis scalable for big data using H2O software, using sample data sets and several machine-learning techniques including deep learning, random forests, unsupervised learning and ensemble learning.**  
**A step-by-step solution-based guide to preparing building, training, and deploying high-quality machine learning models with Amazon SageMaker Key FeaturesPerform ML experiments with built-in and custom algorithms in SageMakerExplore proven solutions when working with TensorFlow, PyTorch, Hugging Face Transformers, and scikit-learnUse the different features and capabilities of SageMaker to automate relevant ML processesBook Description Amazon SageMaker is a fully managed machine learning (ML) service that helps data scientists and ML practitioners manage ML experiments. In this book, you'll use the different capabilities and features of Amazon SageMaker to solve relevant data science and ML problems. This step-by-step guide features 80 proven recipes designed to give you the hands-on machine learning experience needed to contribute to real-world experiments and projects. You'll cover the algorithms and techniques that are commonly used when training and deploying NLP, time series forecasting, and computer vision models to solve ML problems. You'll explore various solutions for working with deep learning libraries and frameworks such as TensorFlow, PyTorch, and Hugging Face Transformers in Amazon SageMaker. You'll also learn how to use SageMaker Clarify, SageMaker Model Monitor, SageMaker Debugger, and**

*SageMaker Experiments to debug, manage, and monitor multiple ML experiments and deployments. Moreover, you'll have a better understanding of how SageMaker Feature Store, Autopilot, and Pipelines can meet the specific needs of data science teams. By the end of this book, you'll be able to combine the different solutions you've learned as building blocks to solve real-world ML problems. What you will learn Train and deploy NLP, time series forecasting, and computer vision models to solve different business problems Push the limits of customization in SageMaker using custom container images Use AutoML capabilities with SageMaker Autopilot to create high-quality models Work with effective data analysis and preparation techniques Explore solutions for debugging and managing ML experiments and deployments Deal with bias detection and ML explainability requirements using SageMaker Clarify Automate intermediate and complex deployments and workflows using a variety of solutions Who this book is for This book is for developers, data scientists, and machine learning practitioners interested in using Amazon SageMaker to build, analyze, and deploy machine learning models with 80 step-by-step recipes. All you need is an AWS account to get things running. Prior knowledge of AWS, machine learning, and the Python programming language will help you to grasp the concepts covered in this book more effectively.*

**R Markdown**

**Powerful, Scalable Techniques for Deep Learning and AI**

**Exam LX0-103 and Exam LX0-104**

**Exams LX0-101 and LX0-102**

**Practical Penetration Testing Techniques**

**Practical Machine Learning with H2O**