
Python Pour Le Data Scientist Des Bases Du Langag

When somebody should go to the books stores, search creation by shop, shelf by shelf, it is in reality problematic. This is why we provide the book compilations in this website. It will completely ease you to see guide Python Pour Le Data Scientist Des Bases Du Langag as you such as.

By searching the title, publisher, or authors of guide you in reality want, you can discover them rapidly. In the house, workplace, or perhaps in your method can be every best area within net connections. If you mean to download and install the Python Pour Le Data Scientist Des Bases Du Langag, it is utterly simple then, since currently we extend the partner to purchase and make bargains to download and install Python Pour Le Data Scientist Des Bases Du Langag thus simple!



L'intelligence artificielle, une révolution ? AMACOM
Le best-seller O'Reilly sur la data science avec Python enfin traduit en Français ? Pour de nombreux chercheurs, Python est un outil essentiel en raison de ses bibliothèques pour stocker, manipuler et obtenir un aperçu des données. Ce livre décrit toutes les ressources dont vous pouvez disposer pour mettre en oeuvre vos applications : IPython, NumPy, Pandas, Matplotlib, Scikit-Learn et d'autres outils associés. Les scientifiques en activité et les utilisateurs de données

familiarisés avec la lecture et l'écriture de code Python trouveront avec cette référence complète l'outil idéal pour s'attaquer aux problèmes quotidiens : manipuler, transformer et nettoyer les données ; visualiser différents types de données ; utiliser les données pour créer des modèles statistiques ou d'apprentissage automatique. Ce livre est tout simplement la référence incontournable pour le calcul scientifique en Python. Vous apprendrez à utiliser : IPython et Jupyter : fournissent des environnements de calcul pour les data scientists utilisant Python NumPy : inclut le ndarray pour un stockage et une manipulation efficaces de tableaux de données denses en Python Pandas : comprend le DataFrame pour un stockage et une manipulation efficaces des

données étiquetées / en colonnes en Python Matplotlib: inclut des fonctionnalités pour une gamme flexible de visualisations de données en Python Scikit-Learn: pour des implémentations Python efficaces et propres des algorithmes d'apprentissage automatique les plus importants et les plus établis Python pour la Data science Pour les Nuls National Academies Press

Un livre à la fois théorique et pratique. Cet ouvrage a pour ambition de couvrir un large spectre du domaine des data sciences. Il va plus loin qu'un simple tour sur les algorithmes d'apprentissage automatique et s'attaque aux autres aspects, malheureusement négligés mais fondamentaux pour tout data scientist : - concepts généraux mais poussés, dont la maîtrise est indispensable ; - algorithmes d'apprentissage automatique les plus connus ; - aspects liés à l'exploration des données ; - mesures de performances et d'autres métriques utilisées par les algorithmes ; - différents concepts fondamentaux en mathématiques à connaître pour mieux explorer et comprendre les données ; - notions importantes des big data ; - études de cas pratiques en langage Python. La première partie de ce livre est théorique et adopte un format questions/réponses qui présente plusieurs avantages, dont la possibilité pour le lecteur de lire distinctement chacune des questions pour parfaire son savoir. Un autre point fort de cette structure est qu'elle incite à entrer dans un dialogue. Ainsi, grâce aux questions posées, le lecteur est poussé vers une réflexion où il confronte ses réponses à celles données par le livre. La deuxième partie est pratique et propose deux exemples d'implémentation de modèles d'apprentissage automatique. Vous y trouverez des codes écrits en Python et un aperçu de différents

difficultés que peut rencontrer un spécialiste lors de l'exercice de son métier

Artificial Intelligence with Python
Editions Eyrolles

Vous aussi participez à la révolution qui ramène l'intelligence artificielle au cœur de notre société, grâce aux data scientists. La data science consiste à traduire des problèmes de toute autre nature, en problèmes de modélisation quantitative, résolus par des algorithmes de traitement. Ce livre se présente comme une référence pour tous les développeurs, statisticiens ou chefs de projets ayant à résoudre des problèmes liés à la data science. Au programme : Pourquoi utiliser le machine learning Les différentes versions de Python L'apprentissage non supervisé et le préprocessing Représenter les données Processus de validation Algorithmes, chaînes et pipeline Travailler avec des données de type texte Du prototype à la production

The Python Bible 7 in 1 "O'Reilly Media, Inc."

Build real-world Artificial Intelligence applications with Python to intelligently interact with the world around you About This Book Step into the amazing world of intelligent apps using this comprehensive guide Enter the world of Artificial Intelligence, explore it, and create your own applications Work through simple yet insightful examples that will get you up and running with Artificial Intelligence in no time Who This Book Is For This book is for Python developers who want to build real-world Artificial Intelligence applications. This book is friendly to Python beginners, but being familiar with Python would be useful to play around with the code. It will also be useful for experienced Python programmers who are looking to use Artificial Intelligence techniques in their existing technology stacks. What You Will Learn Realize different classification and regression techniques Understand the concept of clustering and how to use it to automatically segment data See how to

build an intelligent recommender system Understand logic programming and how to use it Build automatic speech recognition systems Understand the basics of heuristic search and genetic programming Develop games using Artificial Intelligence Learn how reinforcement learning works Discover how to build intelligent applications centered on images, text, and time series data See how to use deep learning algorithms and build applications based on it In Detail Artificial Intelligence is becoming increasingly relevant in the modern world where everything is driven by technology and data. It is used extensively across many fields such as search engines, image recognition, robotics, finance, and so on. We will explore various real-world scenarios in this book and you'll learn about various algorithms that can be used to build Artificial Intelligence applications. During the course of this book, you will find out how to make informed decisions about what algorithms to use in a given context. Starting from the basics of Artificial Intelligence, you will learn how to develop various building blocks using different data mining techniques. You will see how to implement different algorithms to get the best possible results, and will understand how to apply them to real-world scenarios. If you want to add an intelligence layer to any application that's based on images, text, stock market, or some other form of data, this exciting book on Artificial Intelligence will definitely be your guide! Style and approach This highly practical book will show you how to implement Artificial Intelligence. The book provides multiple examples enabling you to create smart applications to meet the needs of your organization. In every chapter, we explain an algorithm, implement it, and then build a smart application.

Machine learning avec Python ANNULE BPB Publications

Understand the constructs of the Python programming language and use them to build data science projects Key Features Learn the basics of developing applications with Python and deploy your first data application Take your first steps in Python programming by understanding and using data structures, variables, and loops Delve into Jupyter, NumPy, Pandas, SciPy, and sklearn to explore the data science ecosystem in Python Book Description Python is the most widely used programming language for building data science applications. Complete with step-by-step instructions, this book contains easy-to-follow tutorials to help you learn Python and develop real-world data science projects. The “secret sauce” of the book is its curated list of topics and solutions, put together using a range of real-world projects, covering initial data collection, data analysis, and production. This Python book starts by taking you through the basics of programming, right from variables and data types to classes and functions. You ’ ll learn how to write idiomatic code and test and debug it, and discover how you can create packages or use the range of built-in ones. You ’ ll also be introduced to the extensive ecosystem of Python data science packages, including NumPy, Pandas, scikit-learn, Altair, and Datashader. Furthermore, you ’ ll be able to perform data analysis, train models, and interpret and communicate the results. Finally, you ’ ll get to grips with structuring and scheduling scripts using Luigi and sharing your machine learning models with the world as a microservice. By the end of the book, you ’ ll have learned not only how to implement Python in data science projects, but also how to maintain and design them to meet high programming standards. What you will learn Code in Python using Jupyter and VS Code Explore the basics of coding – loops, variables, functions, and classes Deploy continuous integration with Git, Bash, and DVC Get to grips with Pandas, NumPy, and scikit-learn Perform data visualization with Matplotlib, Altair, and Datashader Create a package out of your code using poetry and test it with PyTest Make your machine learning model accessible to anyone with the web API Who this book is for If you want to learn Python or data science in a fun and engaging way, this book is for you. You ’ ll also find this book

useful if you're a high school student, researcher, analyst, or anyone with little or no coding experience with an interest in the subject and courage to learn, fail, and learn from failing. A basic understanding of how computers work will be useful.

Interpretable Machine Learning with Python First Interactive

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

Mastering Social Media Mining with Python Alisio

Master the essential skills needed to recognize and solve complex problems with machine learning and deep learning. Using real-world examples that leverage the popular Python machine learning ecosystem, this book is your perfect companion for learning the art and science of machine learning to become a successful practitioner.

The concepts, techniques, tools, frameworks, and methodologies used in this book will teach you how to think, design, build, and execute machine learning systems and projects successfully. Practical Machine Learning with Python follows a structured and comprehensive three-tiered approach packed with hands-on examples and code. Part 1 focuses on understanding machine learning concepts and tools. This includes machine learning basics with a broad overview of algorithms, techniques, concepts and applications, followed by a tour of the entire Python machine learning ecosystem. Brief guides for useful machine learning tools, libraries and frameworks are also covered. Part 2 details standard machine learning pipelines, with an emphasis on data processing analysis, feature engineering, and modeling. You will learn how to process, wrangle, summarize and visualize data in its various forms. Feature engineering and selection methodologies will be covered in detail with real-world datasets followed by model building, tuning, interpretation and deployment. Part 3 explores multiple real-world case studies spanning diverse domains and industries like retail, transportation, movies, music, marketing, computer vision and finance. For each case study, you will learn the application of various machine learning techniques and methods. The hands-on examples will help you become familiar with state-of-the-art machine learning tools and techniques and understand what algorithms are best suited for any problem. Practical Machine Learning with Python will empower you to start solving your own

problems with machine learning today!
What You'll Learn Execute end-to-end machine learning projects and systems
Implement hands-on examples with industry standard, open source, robust machine learning tools and frameworks
Review case studies depicting applications of machine learning and deep learning on diverse domains and industries
Apply a wide range of machine learning models including regression, classification, and clustering. Understand and apply the latest models and methodologies from deep learning including CNNs, RNNs, LSTMs and transfer learning. Who This Book Is For IT professionals, analysts, developers, data scientists, engineers, graduate students
Data Science avec Python Walter de Gruyter GmbH & Co KG
D é couvrez Python le langage de pr é dilection de la science des donn é es
La science des donn é es ou data science consiste à extraire des connaissances dans un flot de donn é es. Elle utilise des techniques et des th é ories tir é es de domaines comme les math é matiques et la statistique. Les m é thodes qui s'adaptent aux big data sont particuli è rement int é ressantes dans la science des donn é es. Ce livre est destin é à tous ceux qui travaillent dans le domaine de l'analyse de donn é es et montre comment mettre en oeuvre le langage Python pour dans ce domaine et uassi celui de la statistique. Il traite é galement de Google Colab, un outil qui permet d'crire du code Python dans le cloud. Au programme : Probabilit é s Distributions al é atoirs Tests d'hypoth è ses Regressions Mod è les de pr é diction
Python for Data Analysis CRC Press

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

Python pour le data scientist John Wiley & Sons

Un ouvrage de r é f é rence pour les (futurs) data scientists Les biblioth è ques, les frameworks, les modules et les bo î tes à outils sont parfaits pour faire de la data science. Ils sont aussi un bon moyen de plonger dans cette discipline sans la comprendre. Dans cet ouvrage, vous apprendrez comment fonctionnent les outils et algorithmes les plus fondamentaux de la data science, en les r é alisant à partir de z é ro. Si

vous êtes fort en maths et que vous connaissez la programmation, l'auteur, Joel Grus, vous aidera à vous familiariser avec les maths et les statistiques qui sont au cœur de la data science et avec les compétences informatiques indispensables pour démarrer comme data scientist. La profusion des données d'aujourd'hui contient les réponses aux questions que personne n'a encore pensé à poser. Ce livre vous enseigne comment obtenir ces réponses. La deuxième édition, revue et augmentée, de cet ouvrage comporte des codes et exemples associés entièrement réécrits en Python 3.6 et intègre de nouveaux chapitres sur l'apprentissage profond (deep learning), les statistiques et le traitement en langage naturel. Suivez un cours accéléré de Python. Apprenez les fondamentaux de l'algèbre linéaire, des statistiques et des probabilités, et comprenez comment et quand les utiliser en data science. Collectez, explorez, nettoyez, bricolez et manipulez les données. Plongez dans les bases de l'apprentissage automatique. Implémentez des modèles comme les k plus proches voisins, la classification naïve bayésienne, les régressions linéaire ou logistique, les arbres de décision, les réseaux neuronaux et le clustering. Explorez les systèmes de recommandation, le traitement du langage naturel, l'analyse de réseau, MapReduce et les bases de données. À qui s'adresse cet ouvrage ? Aux développeurs, statisticiens, étudiants et chefs de projet ayant à résoudre des problèmes de data science. Aux data scientists, mais aussi à toute personne curieuse d'avoir une vue d'ensemble de l'état de l'art de ce

métier du futur.

Think Bigger Packt Publishing Ltd
Understand the key aspects and challenges of machine learning interpretability, learn how to overcome them with interpretation methods, and leverage them to build fairer, safer, and more reliable models
Key Features
Learn how to extract easy-to-understand insights from any machine learning model
Become well-versed with interpretability techniques to build fairer, safer, and more reliable models
Mitigate risks in AI systems before they have broader implications by learning how to debug black-box models
Book Description
Do you want to understand your models and mitigate risks associated with poor predictions using machine learning (ML) interpretation? Interpretable Machine Learning with Python can help you work effectively with ML models. The first section of the book is a beginner's guide to interpretability, covering its relevance in business and exploring its key aspects and challenges. You'll focus on how white-box models work, compare them to black-box and glass-box models, and examine their trade-off. The second section will get you up to speed with a vast array of interpretation methods, also known as Explainable AI (XAI) methods, and how to apply them to different use cases, be it for classification or regression, for tabular, time-series, image or text. In addition to the

step-by-step code, the book also helps the reader to interpret model outcomes using examples. In the third section, you'll get hands-on with tuning models and training data for interpretability by reducing complexity, mitigating bias, placing guardrails, and enhancing reliability. The methods you'll explore here range from state-of-the-art feature selection and dataset debiasing methods to monotonic constraints and adversarial retraining. By the end of this book, you'll be able to understand ML models better and enhance them through interpretability tuning. What you will learn

Recognize the importance of interpretability in business
Study models that are intrinsically interpretable such as linear models, decision trees, and Naïve Bayes
Become well-versed in interpreting models with model-agnostic methods
Visualize how an image classifier works and what it learns
Understand how to mitigate the influence of bias in datasets
Discover how to make models more reliable with adversarial robustness
Use monotonic constraints to make fairer and safer models

Who this book is for
This book is for data scientists, machine learning developers, and data stewards who have an increasingly critical responsibility to explain how the AI systems they develop work, their impact on decision making, and how they identify and manage bias. Working knowledge of machine learning and

the Python programming language is expected.

Data Analysis with Open Source Tools First Interactive

Entrez de plain-pied dans le monde fascinant la data science Vous aussi participez à la révolution qui ramène l'intelligence artificielle au coeur de notre société, grâce aux data scientists. La data science consiste à traduire des problèmes de toute autre nature, en problèmes de modélisation quantitative, résolus par des algorithmes de traitement. Ce livre se présente comme une référence pour tous les développeurs, statisticiens ou chefs de projets ayant à résoudre des problèmes liés à la data science.

Au programme : Pourquoi utiliser le machine learning
Les différences versions de Python
L'apprentissage non supervisé et le préprocessing
Représenter les données
Processus de validation
Algorithmes, chaînes et pipeline
Travailler avec des données de type texte
Du prototype à la production

Data Science for Supply Chain Forecasting First Interactive

Digging into data does not have to be painful. With Data Wrangling Using Python, you'll learn how to clean and analyze data, create compelling stories, and scale that data as necessary. There are awesome discoveries to be made in unassuming datasets and stories to be told. You don't have to be a programmer to tell them. What you need is to understand the context of the data and to know a few of the techniques found in this book. You'll learn enough Python to be empowered to engage with your data, through a series of examples that grow in complexity throughout the book.

Python pour la data science pour les nuls
Eyrolles

Data science is emerging as a field that is

revolutionizing science and industries alike. Work across nearly all domains is becoming more data driven, affecting both the jobs that are available and the skills that are required. As more data and ways of analyzing them become available, more aspects of the economy, society, and daily life will become dependent on data. It is imperative that educators, administrators, and students begin today to consider how to best prepare for and keep pace with this data-driven era of tomorrow.

Undergraduate teaching, in particular, offers a critical link in offering more data science exposure to students and expanding the supply of data science talent. Data Science for Undergraduates: Opportunities and Options offers a vision for the emerging discipline of data science at the undergraduate level. This report outlines some considerations and approaches for academic institutions and others in the broader data science communities to help guide the ongoing transformation of this field.

Python pour la Data Science BoD - Books on Demand

Python est devenu en quelques années un langage majeur dans l'univers des applications centrées sur le traitement des données, et plus particulièrement des gros volumes de données (big data). Cet ouvrage servira de guide à tous ceux qui s'intéressent à l'utilisation de Python pour le travail sur les données et l'automatisation de certaines tâches (data science). Il met l'accent sur la préparation et la mise en forme des données qui sont essentielles dans la qualité du résultat et qui constituent aujourd'hui une part importante du travail du data scientist. L'ensemble des exemples et des exercices présentés dans cet ouvrage sont

disponibles sous forme de Notebooks Jupyter. Ils sont accessibles directement sur GitHub dans le répertoire dédié à l'ouvrage ou en téléchargement sur le site Dunod. Cette 2e édition comportent de très nombreuses mises à jour sur les évolutions récentes du langage Python, sur les "packages" utilisés en data science, et sur d'autres logiciels associés comme TensorFlow et Keras.

Deep Learning with Python First Interactive

Le best-seller O'Reilly sur la data science avec Python enfin traduit en Français ? Pour de nombreux chercheurs, Python est un outil essentiel en raison de ses bibliothèques pour stocker, manipuler et obtenir un aperçu des données. Ce livre décrit toutes les ressources dont vous pouvez disposer pour mettre en œuvre vos applications : IPython, NumPy, Pandas, Matplotlib, Scikit-Learn et d'autres outils associés. Les scientifiques en activité et les utilisateurs de données familiarisés avec la lecture et l'écriture de code Python trouveront avec cette référence complète l'outil idéal pour s'attaquer aux problèmes quotidiens : manipuler, transformer et nettoyer les données ; visualiser différents types de données ; utiliser les données pour créer des modèles statistiques ou d'apprentissage automatique. Ce livre est tout simplement la référence incontournable pour le calcul scientifique en Python. Vous apprendrez à utiliser : IPython et Jupyter : fournissent des environnements de calcul pour les data

scientists utilisant Python NumPy : inclut le ndarray pour un stockage et une manipulation efficaces de tableaux de données denses en Python Pandas : comprend le DataFrame pour un stockage et une manipulation efficaces des données étiquetées / en colonnes en Python Matplotlib : inclut des fonctionnalités pour une gamme flexible de visualisations de données en Python Scikit-Learn : pour des implémentations Python efficaces et propres des algorithmes d'apprentissage automatique les plus importants et les plus établis

Data science pour l'agriculture et l'environnement Packt Publishing Ltd

Vous êtes à la recherche d'un cours accéléré de Python pour la science des données et vous voulez créer facilement votre premier projet en un rien de temps? Vous êtes constamment à la recherche d'informations sur les réseaux sociaux (comme les groupes FB) et vous ne savez pas par où commencer avec la programmation Python? Si c'est le cas, lisez la suite! Python est souvent utilisé en science des données aujourd'hui car c'est un langage de programmation mature qui possède de d'excellentes propriétés pour les programmeurs débutants. Parmi les plus notables de ces propriétés, citons le mot de passe facile à lire, la suppression des délimiteurs optionnels, l'écriture dynamique et l'utilisation de la mémoire dynamique. L'amélioration et la recherche très utile dans le monde de l'informatique et de la technologie ont accru l'importance de ses concepts les plus fondamentaux et essentiels sous mille aspects. Cette notion de principe est ce que nous

appelons continuellement les données, et ces données sont la seule chose qui ouvre la voie à tout dans le monde. Les plus grandes organisations et entreprises du monde ont bâti leur création et leurs philosophies et déterminent une part unique de leur rémunération grâce aux données. La valeur et l'importance des données peuvent être comprises avec la simple certitude qu'un centre de stockage/distribution de données légitime est un million de fois plus rentable que la mine d'or pur du monde avancé. Cependant, l'apprentissage de toutes les compétences requises pour maîtriser la science des données et l'apprentissage automatique pourrait certainement être un défi. MAIS NE VOUS INQUIÉTEZ PAS: dans ce guide complet, nous avons condensé toutes les connaissances dont vous avez besoin d'une manière simple et pratique. Dans ce livre, vous êtes prêt à découvrir: - Comment faire vos premiers pas dans le monde de "Python". Je vous expliquerai, avec des images faciles à suivre, comment installer Python sur les systèmes Mac OS X, Windows et Linux. - Comment mettre en place facilement votre premier projet de Data Science à partir de zéro avec Python en moins de 7 jours. - Des codes et des exercices pratiques pour utiliser Python. Je vous expliquerai le processus étape par étape pour créer des jeux comme: "magic 8 ball" et "hangman game". - Comment fonctionnent les algorithmes de régression utilisés en science des données et quels sont les meilleurs trucs et astuces pour travailler avec eux. - Comment la bibliothèque Scikit-Learn est utilisée dans le

développement d'un algorithme d'apprentissage automatique. - Et bien d'autres choses encore! Même si vous êtes encore un débutant qui se bat pour savoir comment démarrer des projets avec Python, ce livre vous donnera certainement les bonnes informations pour faire passer vos compétences en programmation au niveau supérieur. Procurez-vous votre propre exemplaire dès aujourd'hui en cliquant sur le bouton **ACHETER MAINTENANT** en haut de la page! [JavaScript for Data Science](#) Packt Publishing Ltd

This new edited volume consists of a collection of original articles written by leading financial economists and industry experts in the area of machine learning for asset management. The chapters introduce the reader to some of the latest research developments in the area of equity, multi-asset and factor investing. Each chapter deals with new methods for return and risk forecasting, stock selection, portfolio construction, performance attribution and transaction costs modeling. This volume will be of great help to portfolio managers, asset owners and consultants, as well as academics and students who want to improve their knowledge of machine learning in asset management.

Python Programmation Dunod
Using data science in order to solve a problem requires a scientific mindset more than coding skills. Data Science for Supply Chain Forecasting, Second Edition contends that a true scientific method which includes experimentation, observation, and constant questioning must be applied to supply chains to achieve excellence

in demand forecasting. This second edition adds more than 45 percent extra content with four new chapters including an introduction to neural networks and the forecast value added framework. Part I focuses on statistical "traditional" models, Part II, on machine learning, and the all-new Part III discusses demand forecasting process management. The various chapters focus on both forecast models and new concepts such as metrics, underfitting, overfitting, outliers, feature optimization, and external demand drivers. The book is replete with do-it-yourself sections with implementations provided in Python (and Excel for the statistical models) to show the readers how to apply these models themselves. This hands-on book, covering the entire range of forecasting—from the basics all the way to leading-edge models—will benefit supply chain practitioners, forecasters, and analysts looking to go the extra mile with demand forecasting.

Building Data Science Applications with FastAPI "O'Reilly Media, Inc."
Collecting data is relatively easy, but turning raw information into something useful requires that you know how to extract precisely what you need. With this insightful book, intermediate to experienced programmers interested in data analysis will learn techniques for working with data in a business environment. You'll learn how to look at data to discover what it contains, how to capture those ideas in conceptual models, and then feed your understanding back into the organization through business plans, metrics dashboards,

and other applications. Along the way, you'll experiment with concepts through hands-on workshops at the end of each chapter. Above all, you'll learn how to think about the results you want to achieve -- rather than rely on tools to think for you. Use graphics to describe data with one, two, or dozens of variables Develop conceptual models using back-of-the-envelope calculations, as well as scaling and probability arguments Mine data with computationally intensive methods such as simulation and clustering Make your conclusions understandable through reports, dashboards, and other metrics programs Understand financial calculations, including the time-value of money Use dimensionality reduction techniques or predictive analytics to conquer challenging data analysis situations Become familiar with different open source programming environments for data analysis "Finally, a concise reference for understanding how to conquer piles of data."--Austin King, Senior Web Developer, Mozilla "An indispensable text for aspiring data scientists."--Michael E. Driscoll, CEO/Founder, Dataspora