# Hands On Data Science With Sql Server 2017 Perfor

Big Data Science & Analytics
Data Science on AWS
Hands-On Data Science with Anaconda
Python Data Science
End-to-End Data Science with SAS
Hands-On Data Science with R
Data Analysis with Open Source Tools
Hands-On Big Data Analytics with PySpark
Practical Data Science with Python
Introduction to Data Science
A Hands-On Introduction to Data Science
Hands-on Scikit-Learn for Machine Learning Applications
Beginning Data Science with Python and Jupyter
Data Science Projects with Python
Practical Big Data Analytics

Hands on Data Science for Biologists Using Python
Python for Data Science
R for Data Science
Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow
Getting Started with Data Science
Grokking Deep Learning
Hands-On Data Science for Marketing
Introduction to Data Science
Data Science Projects with Python
Hands-On Data Science and Python Machine Learning
Hands-On Exploratory Data Analysis with R
Data Science For Dummies
Guide to Intelligent Data Science
Python for Data Science
Data Science at the Command Line
Data Science on the Google Cloud Platform
Data Science Using Python and R
Hands-On Machine Learning with R
Data Science from Scratch
Hands-On Data Analysis with Pandas

Network Data Analytics
Hands-On Data Preprocessing in Python
Python Data Science Handbook
Foundations of Data Science
Hands-On Data Analysis with Pandas



## ASHTYN BLANKENSHIP

Big Data Science & Analytics Packt Publishing Ltd
Aspiring data science professionals can learn the Scikit-Learn library along with the fundamentals of machine learning with this book. The book combines the Anaconda Python distribution with the popular Scikit-Learn library to demonstrate a wide range of supervised and unsupervised machine learning algorithms. Care is taken to walk you through the principles of machine learning through clear examples written in Python that you can try out and experiment with at home on your own machine. All applied math and programming skills required to master the content are covered in this book. In-depth knowledge of object-oriented programming is not required as working and complete examples are provided and explained. Coding examples are in-depth and complex when necessary. They are also concise, accurate, and

complete, and complement the machine learning concepts introduced. Working the examples helps to build the skills necessary to understand and apply complex machine learning algorithms. Hands-on Scikit-Learn for Machine Learning Applications is an excellent starting point for those pursuing a career in machine learning. Students of this book will learn the fundamentals that are a prerequisite to competency. Readers will be exposed to the Anaconda distribution of Python that is designed specifically for data science professionals, and will build skills in the popular Scikit-Learn library that underlies many machine learning applications in the world of Python. What You'll LearnWork with simple and complex datasets common to Scikit-Learn Manipulate data into vectors and matrices for algorithmic processing Become familiar with the Anaconda distribution used in data scienceApply machine learning with Classifiers, Regressors, and Dimensionality Reduction Tune algorithms and find the best algorithms for each dataset Load data from and save to CSV, JSON, Numpy, and Pandas formats Who This Book Is For The aspiring data scientist yearning to break into machine learning through mastering the underlying fundamentals that are sometimes skipped over in the rush to be productive. Some knowledge of object-oriented programming and very basic applied linear algebra will make learning easier, although anyone can benefit from this book.

*Data Science on AWS* CRC Press
An introductory textbook offering a low barrier entry to data science; the hands-on approach will appeal to students from a range of disciplines.
Hands-On Data Science with Anaconda Packt Publishing Ltd
Data science libraries, frameworks, modules, and toolkits are great for doing data science, but they're also a good way to dive into the discipline without actually understanding data science. In this book, you'll learn how many of the most fundamental data science tools and algorithms work by implementing them from scratch. If you have an aptitude for mathematics and some programming skills, author Joel Grus will help you get comfortable with the math and statistics at the core of data science, and with hacking skills you need to get started as a data scientist. Today's messy glut of data holds answers to questions no one's even thought to ask. This book provides you with the know-how to dig those answers out. Get a crash course in Python Learn the basics of linear algebra, statistics, and probability—and understand how and when they're used in data science Collect, explore, clean, munge, and manipulate data Dive into the fundamentals of machine learning Implement models such as k-nearest Neighbors, Naive Bayes, linear and logistic regression, decision trees, neural networks, and clustering Explore recommender systems, natural language processing, network analysis, MapReduce, and databases

*Python Data Science* "O'Reilly Media, Inc."
Learn exploratory data analysis concepts using powerful R packages to enhance your R data analysis skills Key FeaturesSpeed up your data analysis projects using powerful R packages and techniquesCreate multiple hands-on data analysis projects using real-world dataDiscover and practice graphical exploratory analysis techniques across domainsBook Description Hands-On Exploratory Data Analysis with R will help you build not just a foundation but also expertise in the elementary ways to analyze data. You will learn how to understand your data and summarize its main characteristics. You'll also uncover the structure of your data, and you'll learn graphical and numerical techniques using the R language. This book covers the entire exploratory data analysis (EDA) process—data collection, generating statistics, distribution, and invalidating the hypothesis. As you progress through the book, you will learn how to set up a data analysis environment with tools such as ggplot2, knitr, and R Markdown, using tools such as DOE Scatter Plot and SML2010 for multifactor, optimization, and regression data problems. By the end of this book, you will be able to successfully carry out a preliminary investigation on any dataset, identify hidden insights, and present your results in a business context. What you will learnLearn powerful R techniques to speed up your data analysis projectsImport, clean, and explore data using powerful R

packagesPractice graphical exploratory analysis techniquesCreate informative data analysis reports using ggplot2Identify and clean missing and erroneous dataExplore data analysis techniques to analyze multi-factor datasetsWho this book is for Hands-On Exploratory Data Analysis with R is for data enthusiasts who want to build a strong foundation for data analysis. If you are a data analyst, data engineer, software engineer, or product manager, this book will sharpen your skills in the complete workflow of exploratory data analysis.

**End-to-End Data Science with SAS**

Packt Publishing Ltd

Through a series of recent breakthroughs, deep learning has boosted the entire field of machine learning. Now, even programmers who know close to nothing about this technology can use simple, efficient tools to implement programs capable of learning from data. This practical book shows you how. By using concrete examples, minimal theory, and two production-ready Python frameworks—Scikit-Learn and TensorFlow—author Aurélien Géron helps you gain an intuitive understanding of the concepts and tools for building intelligent systems. You'll learn a range of techniques, starting with simple linear regression and progressing to deep neural networks. With exercises in each chapter to help you apply what you've learned, all you need is programming experience to get started. Explore the machine learning landscape,

particularly neural nets Use Scikit-Learn to track an example machine-learning project end-to-end Explore several training models, including support vector machines, decision trees, random forests, and ensemble methods Use the TensorFlow library to build and train neural nets Dive into neural net architectures, including convolutional nets, recurrent nets, and deep reinforcement learning Learn techniques for training and scaling deep neural nets
Hands-On Data Science with R SAS Institute
Use PySpark to easily crush messy data at-scale and discover proven techniques to create testable, immutable, and easily parallelizable Spark jobs Key FeaturesWork with large amounts of agile data using distributed datasets and in-memory cachingSource data from all popular data hosting platforms, such as HDFS, Hive, JSON, and S3Employ the easy-to-use PySpark API to deploy big data Analytics for productionBook Description Apache Spark is an open source parallel-processing framework that has been around for quite some time now. One of the many uses of Apache Spark is for data analytics applications across clustered computers. In this book, you will not only learn how to use Spark and the Python API to create high-performance analytics with big data, but also discover techniques for testing, immunizing, and parallelizing Spark jobs. You will learn how to source data from all popular data hosting platforms, including HDFS, Hive, JSON, and S3, and deal with large datasets

with PySpark to gain practical big data experience. This book will help you work on prototypes on local machines and subsequently go on to handle messy data in production and at scale. This book covers installing and setting up PySpark, RDD operations, big data cleaning and wrangling, and aggregating and summarizing data into useful reports. You will also learn how to implement some practical and proven techniques to improve certain aspects of programming and administration in Apache Spark. By the end of the book, you will be able to build big data analytical solutions using the various PySpark offerings and also optimize them effectively. What you will learnGet practical big data experience while working on messy datasetsAnalyze patterns with Spark SQL to improve your business intelligenceUse PySpark's interactive shell to speed up development timeCreate highly concurrent Spark programs by leveraging immutabilityDiscover ways to avoid the most expensive operation in the Spark API: the shuffle operationRe-design your jobs to use reduceByKey instead of groupByCreate robust processing pipelines by testing Apache Spark jobsWho this book is for This book is for developers, data scientists, business analysts, or anyone who needs to reliably analyze large amounts of large-scale, real-world data. Whether you're tasked with creating your company's business intelligence function or creating great data platforms for your machine learning models, or are looking

to use code to magnify the impact of your business, this book is for you.

Data Analysis with Open Source Tools

John Wiley & Sons

Getting started with data science doesn't have to be an uphill battle. This step-by-step guide is ideal for beginners who know a little Python and are looking for a quick, fast-paced introduction. Key Features Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts like SVM, KNN classifiers and Random Forests Discover how you can use web scraping to gather and parse your own bespoke datasets Book Description Get to grips with the skills you need for entry-level data science in this hands-on Python and Jupyter course. You'll learn about some of the most commonly used libraries that are part of the Anaconda distribution, and then explore machine learning models with real datasets to give you the skills and exposure you need for the real world. We'll finish up by showing you how easy it can be to scrape and gather your own data from the open web, so that you can apply your new skills in an actionable context. What you will learn Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts like SVM, KNN classifiers, and Random Forests Plan a machine learning classification strategy and train classification, models Use validation curves and dimensionality reduction to tune and enhance your models Discover how you can use web

scraping to gather and parse your own bespoke datasets Scrape tabular data from web pages and transform them into Pandas DataFrames Create interactive, web-friendly visualizations to clearly communicate your findings Who this book is for This book is ideal for professionals with a variety of job descriptions across large range of industries, given the rising popularity and accessibility of data science. You'll need some prior experience with Python, with any prior work with libraries like Pandas, Matplotlib and Pandas providing you a useful head start.

*Hands-On Big Data Analytics with PySpark* Packt Publishing Ltd

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

analysis Program—learn powerful R tools for solving data problems with greater clarity and ease Explore—examine your data, generate hypotheses, and quickly test them Model—provide a low-dimensional summary that captures true "signals" in your dataset Communicate—learn R Markdown for integrating prose, code, and results

Practical Data Science with Python IBM Press

This hands-on guide demonstrates how the flexibility of the command line can help you become a more efficient and productive data scientist. You'll learn how to combine small, yet powerful, command-line tools to quickly obtain, scrub, explore, and model your data. To get you started—whether you're on Windows, OS X, or Linux—author Jeroen Janssens introduces the Data Science Toolbox, an easy-to-install virtual environment packed with over 80 command-line tools. Discover why the command line is an agile, scalable, and extensible technology. Even if you're already comfortable processing data with, say, Python or R, you'll greatly improve your data science workflow by also leveraging the power of the command line. Obtain data from websites, APIs, databases, and spreadsheets Perform scrub operations on plain text, CSV, HTML/XML, and JSON Explore data, compute descriptive statistics, and create visualizations Manage your data science workflow using Drake Create reusable tools from one-liners and existing Python or R code Parallelize and distribute data-intensive

pipelines using GNU Parallel Model data with dimensionality reduction, clustering, regression, and classification algorithms *Introduction to Data Science* Vpt Hands-on Data Science for Biologists using Python has been conceptualized to address the massive data handling needs of modern-day biologists. With the advent of high throughput technologies and consequent availability of omics data, biological science has become a data-intensive field. This hands-on textbook has been written with the inception of easing data analysis by providing an interactive, problem-based instructional approach in Python programming language. The book starts with an introduction to Python and steadily delves into scrupulous techniques of data handling, preprocessing, and visualization. The book concludes with machine learning algorithms and their applications in biological data science. Each topic has an intuitive explanation of concepts and is accompanied with biological examples. Features of this book: The book contains standard templates for data analysis using Python, suitable for beginners as well as advanced learners. This book shows working implementations of data handling and machine learning algorithms using real-life biological datasets and problems, such as gene expression analysis; disease prediction; image recognition; SNP association with phenotypes and diseases. Considering the importance of visualization for data interpretation, especially in biological systems, there is

a dedicated chapter for the ease of data visualization and plotting. Every chapter is designed to be interactive and is accompanied with Jupyter notebook to prompt readers to practice in their local systems. Other avant-garde component of the book is the inclusion of a machine learning project, wherein various machine learning algorithms are applied for the identification of genes associated with age-related disorders. A systematic understanding of data analysis steps has always been an important element for biological research. This book is a readily accessible resource that can be used as a handbook for data analysis, as well as a platter of standard code templates for building models.

A Hands-On Introduction to Data Science

Simon and Schuster

Learn how easy it is to apply sophisticated statistical and machine learning methods to real-world problems when you build on top of the Google Cloud Platform (GCP). This hands-on guide shows developers entering the data science field how to implement an end-to-end data pipeline, using statistical and machine learning methods and tools on GCP. Through the course of the book, you'll work through a sample business decision by employing a variety of data science approaches. Follow along by implementing these statistical and machine learning solutions in your own project on GCP, and discover how this platform provides a transformative and more collaborative way of doing data science. You'll learn how to: Automate and schedule data ingest, using an App

Engine application Create and populate a dashboard in Google Data Studio Build a real-time analysis pipeline to carry out streaming analytics Conduct interactive data exploration with Google BigQuery Create a Bayesian model on a Cloud Dataproc cluster Build a logistic regression machine-learning model with Spark Compute time-aggregate features with a Cloud Dataflow pipeline Create a high-performing prediction model with TensorFlow Use your deployed model as a microservice you can access from both batch and real-time pipelines

Hands-on Scikit-Learn for Machine Learning Applications Packt Publishing Ltd

Gain hands-on experience of Python programming with industry-standard machine learning techniques using pandas, scikit-learn, and XGBoost Key FeaturesThink critically about data and use it to form and test a hypothesisChoose an appropriate machine learning model and train it on your dataCommunicate data-driven insights with confidence and clarityBook Description If data is the new oil, then machine learning is the drill. As companies gain access to ever-increasing quantities of raw data, the ability to deliver state-of-the-art predictive models that support business decision-making becomes more and more valuable. In this book, you'll work on an end-to-end project based around a realistic data set and split up into bite-sized practical exercises. This creates a case-study approach that simulates the working conditions you'll experience in

real-world data science projects. You'll learn how to use key Python packages, including pandas, Matplotlib, and scikit-learn, and master the process of data exploration and data processing, before moving on to fitting, evaluating, and tuning algorithms such as regularized logistic regression and random forest. Now in its second edition, this book will take you through the end-to-end process of exploring data and delivering machine learning models. Updated for 2021, this edition includes brand new content on XGBoost, SHAP values, algorithmic fairness, and the ethical concerns of deploying a model in the real world. By the end of this data science book, you'll have the skills, understanding, and confidence to build your own machine learning models and gain insights from real data. What you will learnLoad, explore, and process data using the pandas Python packageUse Matplotlib to create compelling data visualizationsImplement predictive machine learning models with scikit-learnUse lasso and ridge regression to reduce model overfittingEvaluate random forest and logistic regression model performanceDeliver business insights by presenting clear, convincing conclusionsWho this book is for Data Science Projects with Python – Second Edition is for anyone who wants to get started with data science and machine learning. If you're keen to advance your career by using data analysis and predictive modeling to generate business insights, then this book is the perfect place to begin. To quickly grasp

the concepts covered, it is recommended that you have basic experience of programming with Python or another similar language, and a general interest in statistics.

Beginning Data Science with Python and Jupyter Packt Publishing Ltd

For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the Python Data Science Handbook do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you'll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine

learning algorithms
Data Science Projects with Python
"O'Reilly Media, Inc."
Get to grips with pandas—a versatile and high-performance Python library for data manipulation, analysis, and discovery Key FeaturesPerform efficient data analysis and manipulation tasks using pandasApply pandas to different real-world domains using step-by-step demonstrationsGet accustomed to using pandas as an effective data exploration toolBook Description Data analysis has become a necessary skill in a variety of positions where knowing how to work with data and extract insights can generate significant value. Hands-On Data Analysis with Pandas will show you how to analyze your data, get started with machine learning, and work effectively with Python libraries often used for data science, such as pandas, NumPy, matplotlib, seaborn, and scikit-learn. Using real-world datasets, you will learn how to use the powerful pandas library to perform data wrangling to reshape, clean, and aggregate your data. Then, you will learn how to conduct exploratory data analysis by calculating summary statistics and visualizing the data to find patterns. In the concluding chapters, you will explore some applications of anomaly detection, regression, clustering, and classification, using scikit-learn, to make predictions based on past data. By the end of this book, you will be equipped with the skills you need to use pandas to ensure the veracity of your data, visualize it for effective decision-making, and reliably

reproduce analyses across multiple datasets. What you will learnUnderstand how data analysts and scientists gather and analyze dataPerform data analysis and data wrangling in PythonCombine, group, and aggregate data from multiple sourcesCreate data visualizations with pandas, matplotlib, and seabornApply machine learning (ML) algorithms to identify patterns and make predictionsUse Python data science libraries to analyze real-world datasetsUse pandas to solve common data representation and analysis problemsBuild Python scripts, modules, and packages for reusable analysis codeWho this book is for This book is for data analysts, data science beginners, and Python developers who want to explore each stage of data analysis and scientific computing using a wide range of datasets. You will also find this book useful if you are a data scientist who is looking to implement pandas in machine learning. Working knowledge of Python programming language will be beneficial.

Practical Big Data Analytics Springer Nature

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 asscaling 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

Hands on Data Science for Biologists Using Python "O'Reilly Media, Inc." Monetize your company’s data and data science expertise without spending a fortune on hiring independent strategy consultants to help What if there was one simple, clear process for ensuring

that all your company's data science projects achieve a high a return on investment? What if you could validate your ideas for future data science projects, and select the one idea that's most prime for achieving profitability while also moving your company closer to its business vision? There is. Industry-acclaimed data science consultant, Lillian Pierson, shares her proprietary STAR Framework – A simple, proven process for leading profit-forming data science projects. Not sure what data science is yet? Don't worry! Parts 1 and 2 of Data Science For Dummies will get all the bases covered for you. And if you're already a data science expert? Then you really won't want to miss the data science strategy and data monetization gems that are shared in Part 3 onward throughout this book. Data Science For Dummies demonstrates: The only process you'll ever need to lead profitable data science projects Secret, reverse-engineered data monetization tactics that no one's talking about The shocking truth about how simple natural language processing can be How to beat the crowd of data professionals by cultivating your own unique blend of data science expertise Whether you're new to the data science field or already a decade in, you're sure to learn something new and incredibly valuable from Data Science For Dummies. Discover how to generate massive business wins from your company's data by picking up your copy today.

**Python for Data Science** Packt

Publishing Ltd
A hands-on guide for professionals to perform various data science tasks in R Key FeaturesExplore the popular R packages for data scienceUse R for efficient data mining, text analytics and feature engineeringBecome a thorough data science professional with the help of hands-on examples and use-cases in RBook Description R is the most widely used programming language, and when used in association with data science, this powerful combination will solve the complexities involved with unstructured datasets in the real world. This book covers the entire data science ecosystem for aspiring data scientists, right from zero to a level where you are confident enough to get hands-on with real-world data science problems. The book starts with an introduction to data science and introduces readers to popular R libraries for executing data science routine tasks. This book covers all the important processes in data science such as data gathering, cleaning data, and then uncovering patterns from it. You will explore algorithms such as machine learning algorithms, predictive analytical models, and finally deep learning algorithms. You will learn to run the most powerful visualization packages available in R so as to ensure that you can easily derive insights from your data. Towards the end, you will also learn how to integrate R with Spark and Hadoop and perform large-scale data analytics without much complexity. What you will learnUnderstand the R programming language and its

ecosystem of packages for data scienceObtain and clean your data before processingMaster essential exploratory techniques for summarizing dataExamine various machine learning prediction, modelsExplore the H2O analytics platform in R for deep learningApply data mining techniques to available datasetsWork with interactive visualization packages in RIntegrate R with Spark and Hadoop for large-scale data analyticsWho this book is for If you are a budding data scientist keen to learn about the popular pandas library, or a Python developer looking to step into the world of data analysis, this book is the ideal resource you need to get started. Some programming experience in Python will be helpful to get the most out of this course

R for Data Science Packt Publishing Ltd With this practical book, AI and machine learning practitioners will learn how to successfully build and deploy data science projects on Amazon Web Services. The Amazon AI and machine learning stack unifies data science, data engineering, and application development to help level upyour skills. This guide shows you how to build and run pipelines in the cloud, then integrate the results into applications in minutes instead of days. Throughout the book, authors Chris Fregly and Antje Barth demonstrate how to reduce cost and improve performance. Apply the Amazon AI and ML stack to real-world use cases for natural language processing, computer vision, fraud detection, conversational devices, and more Use

automated machine learning to implement a specific subset of use cases with SageMaker Autopilot Dive deep into the complete model development lifecycle for a BERT-based NLP use case including data ingestion, analysis, model training, and deployment Tie everything together into a repeatable machine learning operations pipeline Explore real-time ML, anomaly detection, and streaming analytics on data streams with Amazon Kinesis and Managed Streaming for Apache Kafka Learn security best practices for data science projects and workflows including identity and access management, authentication, authorization, and more

Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow Springer

Introduction to Data Science: Data Analysis and Prediction Algorithms with R introduces concepts and skills that can help you tackle real-world data analysis challenges. It covers concepts from probability, statistical inference, linear regression, and machine learning. It also helps you develop skills such as R programming, data wrangling, data visualization, predictive algorithm building, file organization with UNIX/Linux shell, version control with Git and GitHub, and reproducible document preparation. This book is a textbook for a first course in data science. No previous knowledge of R is necessary, although some experience with programming may be helpful. The book is divided into six parts: R, data visualization, statistics with R, data wrangling, machine learning, and productivity tools. Each

part has several chapters meant to be presented as one lecture. The author uses motivating case studies that realistically mimic a data scientist’s experience. He starts by asking specific questions and answers these through data analysis so concepts are learned as a means to answering the questions. Examples of the case studies included are: US murder rates by state, self-reported student heights, trends in world health and economics, the impact of vaccines on infectious disease rates, the financial crisis of 2007-2008, election forecasting, building a baseball team, image processing of hand-written digits, and movie recommendation systems. The statistical concepts used to answer the case study questions are only briefly introduced, so complementing with a probability and statistics textbook is highly recommended for in-depth understanding of these concepts. If you read and understand the chapters and complete the exercises, you will be prepared to learn the more advanced concepts and skills needed to become an expert.

Getting Started with Data Science

"O'Reilly Media, Inc."

Optimize your marketing strategies through analytics and machine learning Key FeaturesUnderstand how data science drives successful marketing campaignsUse machine learning for better customer engagement, retention, and product recommendationsExtract insights from your data to optimize marketing strategies and increase profitabilityBook Description Regardless

of company size, the adoption of data science and machine learning for marketing has been rising in the industry. With this book, you will learn to implement data science techniques to understand the drivers behind the successes and failures of marketing campaigns. This book is a comprehensive guide to help you understand and predict customer behaviors and create more effectively targeted and personalized marketing strategies. This is a practical guide to performing simple-to-advanced tasks, to extract hidden insights from the data and use them to make smart business decisions. You will understand what drives sales and increases customer engagements for your products. You will learn to implement machine learning to forecast which customers are more likely to engage with the products and have high lifetime value. This book will also show you how to use machine learning techniques to understand different customer segments and recommend the right products for each customer. Apart from learning to gain insights into consumer behavior using exploratory analysis, you will also learn the concept of A/B testing and implement it using Python and R. By the end of this book, you will be experienced enough with various data science and machine learning techniques to run and manage successful marketing campaigns for your business. What you will learnLearn how to compute and visualize marketing KPIs in Python and RMaster what drives successful marketing campaigns with

data scienceUse machine learning to predict customer engagement and lifetime valueMake product recommendations that customers are most likely to buyLearn how to use A/B testing for better marketing decision makingImplement machine learning to understand different customer segmentsWho this book is for If you are a marketing professional, data scientist, engineer, or a student keen to learn how to apply data science to marketing, this book is what you need! It will be beneficial to have some basic knowledge of either Python or R to work through the examples. This book will also be beneficial for beginners as it covers basic-to-advanced data science concepts and applications in marketing with real-life examples.