

Twitter Thread by [Avi Kumar Talaviya](#)

[Avi Kumar Talaviya](#)

[@avikumart](#)



10 FREE online textbooks for data science you must own to excel in solving data problems■■■

A ■↓

1. 66 Days of Data literacy

Authored by: Alex Chisholm

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/bE24Dlqkeu>

1 Welcome

2 We're all data professionals now

3 Resources

3.1 Shared spreadsheets

3.2 General resources

4 Getting data

4.1 A starting point

4.2 Surveys

4.3 Data dictionaries

4.4 Approaching analysis

4.5 Organizational strategy

5 Classifying data

5.1 Data types

5.2 Categorical

5.3 Numeric

6 Exploring data

66 Days of Data (Literacy)

Data Fundamentals and Visualization Techniques

Alex Chisholm, DataKwery

Last Update: 2021-12-20

1 Welcome

It seems that everything is now a data problem. The global pandemic only accelerated this trend with billions of people pushed further into virtual spaces, leaving a digital trail of data behind them

2. A Beginners guide to clean data

Authored by: Benjamin Greve

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/UhUCoEEGVz>

beginners-guide-to-clean-data

A Beginner's Guide to Clean Data

INTRODUCTION

- Foreword
- The value of data
- The intangible nature of data

MISSING DATA

- Missing values
- Missing value patterns
- Missing value representations
- Missing observations
- Truncated exports
- Handling missing values

Powered By GitBook

A Beginner's Guide to Clean Data

Practical advice to spot and avoid data quality problems. - Benjamin Greve

This is a free version of my book "A Beginner's Guide to Clean Data: Practical advice to spot and avoid data quality problems". If you like the content, feel free to [buy this book on Amazon](#) and/or [leave a positive review](#) there.

Summary

This book will help you to become a better data scientist by showing you the things that can go wrong when working with data - particularly low-quality data. A key difference between a junior and a senior data scientist is the awareness of potential pitfalls. The experienced data scientist will expect them, navigate around them and avoid costly iteration cycles. After reading this book, you will be able to spot data quality problems and deal with them before they can break your work, saving yourself a lot of time.

In the past six years of working in data science, I have made all the mistakes described in this book. Every time, it cost me hours, sometimes days to figure out what the problem was and to fix it. This type of iterative work is what data scientists mean when they talk about how they spend most of their time on data preparation. Yet, for some reason, the art of preparing data and ensuring a sufficiently high level of quality is largely ignored by textbooks, university programs, online courses and industry conferences. That's why I felt the need to write this book and share some of my experiences. It is the hands-on advice that I myself wish I had when I started my career as a data

3. Introduction to Business Analytics

Authored by: Adam Fleischhacker

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/9L6NhGBuZU>

- 1 Welcome
- 2 Becoming a Data-Driven Business Analyst
- 3 The Computing Environment
- 4 R: Basic Usage
- 5 R Packages: causact, tidyverse, etc.
- 6 dplyr: Manipulating Data Frames
- 7 dplyr: Data Manipulation For Insight
- 8 ggplot2: Data Visualization Using The Grammar of Graphics
- 9 ggplot2: The Four Stages of Visualization
- 10 Representing Uncertainty
- 11 Joint Distributions Tell You Everything
- 12 Graphical Models Tell Joint Distribution Stories
- 13 Bayesian Inference On Graphical Models
- 14 Generative DAGs As Business and Mathematical Narratives
- 15 Install TensorFlow



[Link to YouTube playlist for videos that accompany each chapter](#)

A Business Analyst's Introduction to Business Analytics

Intro to Bayesian Business Analytics in the R Eco-System (Featuring R, Rstudio, the tidyverse, graphical models, Bayesian inference, causact, and greta's TensorFlow interface from R.)

Adam Fleischhacker

2022-04-26

Chapter 1 Welcome



Figure 1.1: In this book, you will become the business analyst who easily travels between the real-world of business and the theoretical world of mathematics. You will translate real-world scenarios into both mathematical and computational representations that yield actionable insight. You will then take that insight back to the real-world to persuade

4. AI and machine learning for Finance

Authored by: Michela Cameletti

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/ujE0lvbf4s>

2.5 How to install RStudio

3 Lab 1 - 02/03/2022

3.1 Built-in functions

3.2 Objects and vectors

3.3 Random number generation fr...

3.4 Exercises Lab 1

4 Lab 2 - 10/03/2022

4.1 Data import

4.2 Data selection from a data fra...

4.3 Factor

4.4 Install and load a package

4.5 The pipe operator

4.6 dplyr verbs

4.7 Exercises Lab 2

5 Lab 3 - 16/03/2022

5.1 Frequency distribution

5.2 The ggplot2 library


5.3 Start working with the ggplot ...

AI and Machine Learning For Finance 2021/22

Michela Cameletti

2022-05-14

Chapter 1 Welcome! :)



You are reading the lecture notes of the R lectures for the **AI and Machine Learning for Finance**

5. An introduction to statistical learning

Authored by: Gareth James, Daniela Witten, Trevor Hastie, and Rob Tibshirani

Textbook courtesy: [@DataKwery](#)

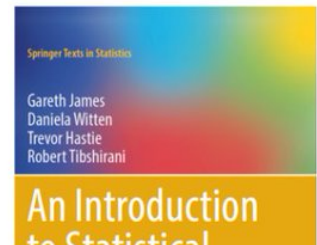
■ <https://t.co/9K5BI91HT>

An Introduction to Statistical Learning

Download the Second Edition

Winner of the 2014 Eric Ziegel award from Technometrics.

As the scale and scope of data collection continue to increase across virtually all fields, statistical learning has become a critical toolkit for anyone who wishes to understand data. *An Introduction to Statistical Learning* provides a broad and less technical treatment of key topics in statistical learning. Each chapter includes an R lab. This book is appropriate for anyone who wishes to use contemporary tools for data analysis.



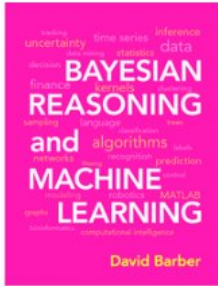
6. Bayesian reading and Machine learning

Authored by: David Barber

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/ORh9GEthj0>

David Barber
Brml :: Online



Online Versions & Errata

The online version differs from the hardcopy in page numbering so please refer to the hardcopy if you wish to cite a particular page.

20 Jun 2020 [🔗](#)

This version corresponds to the published Cambridge University Press version, differing only in minor text details. There are some errata (in magenta) and addenda (in blue) from the published version highlighted using ampersands or plus signs in the margin text.

Please email me at david.barber@ucl.ac.uk if you find an error or have a suggestion.

BRML Homepage
Online version & Errata
Software
Instructors Material
David Barber

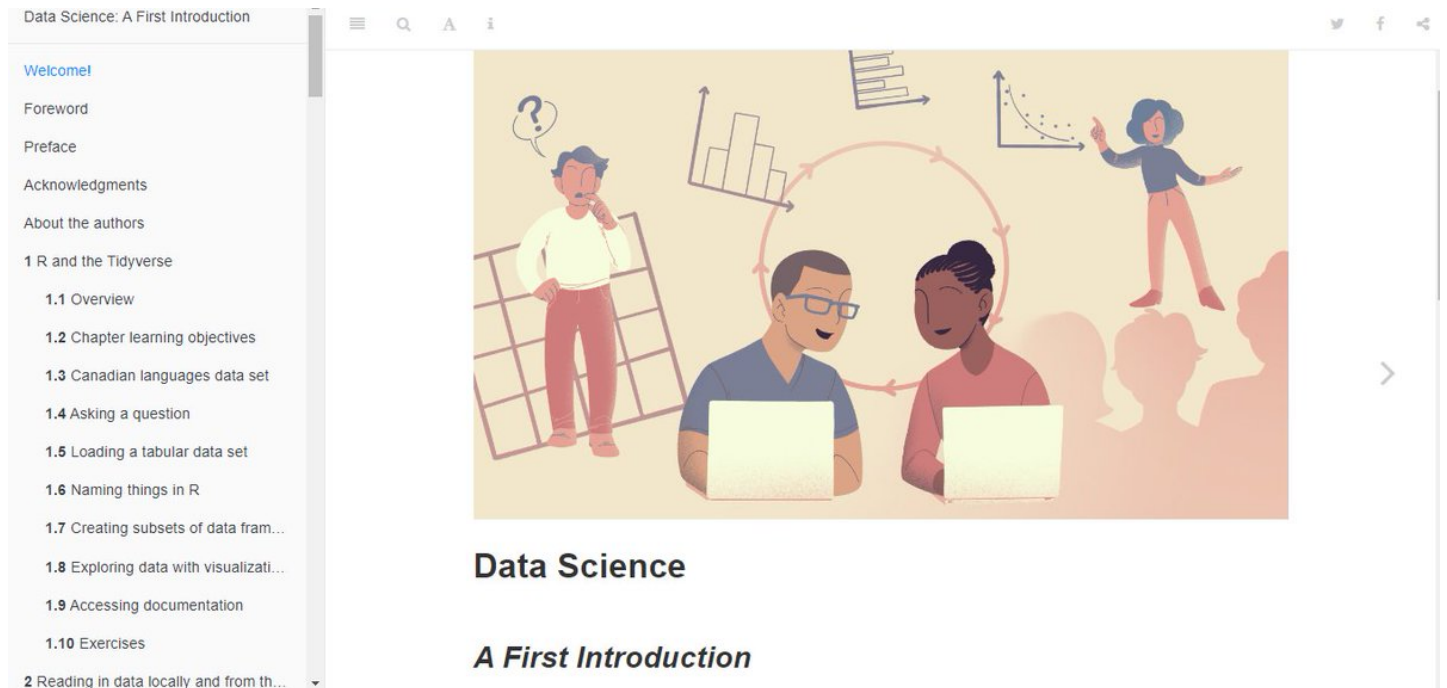
Powered by [PmWiki](#)

7. Data science: First introduction

Authored by: Tiffany Timbers, Trevor Campbell, and Melissa Lee

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/RdCrywxE7D>



Data Science: A First Introduction

- Welcome!
- Foreword
- Preface
- Acknowledgments
- About the authors
- 1 R and the Tidyverse
 - 1.1 Overview
 - 1.2 Chapter learning objectives
 - 1.3 Canadian languages data set
 - 1.4 Asking a question
 - 1.5 Loading a tabular data set
 - 1.6 Naming things in R
 - 1.7 Creating subsets of data fram...
 - 1.8 Exploring data with visualizati...
 - 1.9 Accessing documentation
 - 1.10 Exercises
- 2 Reading in data locally and from th...

Data Science

A First Introduction

8. Python Data Science Handbook

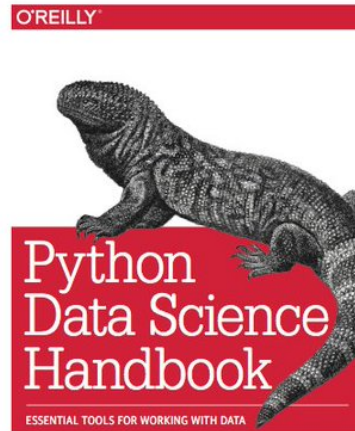
Authored by: Jake VanderPlas

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/J7kwmsLZF4>

Python Data Science Handbook

Jake VanderPlas



9. Hands-On Data Visualization

Authored by: Ilya Ilyankou, Jack Dougherty

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/nrP7SRW9TP>

Hands-On Data Visualization
by Jack Dougherty & Ilya Ilyankou

Preface

- Audience and Overview
- Authors & Acknowledgements
- Our Open-Access Web Edition: W...
- Introduction: Why Data Visualization?
- 1 Choose Tools to Tell Your Story
- 2 Strengthen Your Spreadsheet Skills
- 3 Find and Question Your Data
- 4 Clean Up Messy Data
- 5 Make Meaningful Comparisons
- 6 Chart Your Data
- 7 Map Your Data
- 8 Table Your Data
- 9 Embed On the Web
- 10 Edit and Host Code with GitHub
- 11 Chart.js and Highcharts Templates

Hands-On Data Visualization
Interactive Storytelling from Spreadsheets to Code

Jack Dougherty
Ilya Ilyankou
2022-07-26

Preface

O'REILLY®
Hands-On Data Visualization
Interactive Storytelling from Spreadsheets to Code

10. Deep learning with Python

Authored by: François Chollet

Textbook courtesy: [@DataKwery](#)

■ <https://t.co/tCBm9k00Ve>

table of contents

[🔗 detailed TOC](#)

PART 1: THE FUNDAMENTALS OF DEEP LEARNING

- [1. WHAT IS DEEP LEARNING? ▶](#)
- [2. BEFORE WE START: THE MATHEMATICAL BLOCKS OF NEURAL NETWORKS ▶](#)
- [3. GETTING STARTED WITH NEURAL NETWORKS ▶](#)
- [4. FUNDAMENTALS OF MACHINE LEARNING ▶](#)

PART 2: DEEP LEARNING IN PRACTICE

- [5. DEEP LEARNING FOR COMPUTER VISION ▶](#)
- [6. DEEP LEARNING FOR TEXT AND SEQUENCES ▶](#)
- [7. ADVANCED DEEP LEARNING BEST PRACTICES ▶](#)
- [8. GENERATIVE DEEP LEARNING ▶](#)
- [9. CONCLUSIONS ▶](#)

APPENDIXES

“ Bridges the gap between the hype and a functioning deep-learning system.

Peter Rabinovitch, Akamai

“ The best resource for becoming a master of Keras and deep learning.

Claudio Rodriguez, Cox Media Group

You can read the unrolled version of this thread here: <https://t.co/WGVYIMWDsv>

End of this thread!■

If you've found it informative then do like, RT/QT first tweet, and comment what you think on this■

And Don't forget to follow me at [@avikumart_](#) for more such threads in data science and ML!■

<https://t.co/nMpdqNXpK>

10 FREE online textbooks for data science you must own to excel in solving data problems\U0001f4c8\U0001f680

A \U0001f9f5\U0001f680

— Avi Kumar Talaviya (@avikumart_) [August 2, 2022](#)

I BIG shoutout to [@DataKwery](#) for providing these valuable resources in one place to help the data science community grow!■

Are you looking for the latest courses and resources in data science? then do not forget to sign up to stay updated!■

■<https://t.co/Szdwdp2MIx>

Search Data Science Courses and Resources

Recent Searches:

- Python
- Excel
- SQL
- Tableau
- Power BI
- Machine Learning
- NLP
- Harvard
- Stanford
- MIT
- Google
- Amazon
- Coursera
- DataCamp

