The Art of Data Science

A Guide for Anyone Who Works with Data

Roger D. Peng & Elizabeth Matsui

The Art of Data Science: A Guide for Anyone Who Works with Data

Roger D. Peng , Elizabeth Matsui

The Art of Data Science: A Guide for Anyone Who Works with Data

Roger D. Peng , Elizabeth Matsui

The Art of Data Science: A Guide for Anyone Who Works with Data Roger D. Peng , Elizabeth Matsui This book describes, simply and in general terms, the process of analyzing data. The authors have extensive experience both managing data analysts and conducting their own data analyses, and have carefully observed what produces coherent results and what fails to produce useful insights into data. This book is a distillation of their experience in a format that is applicable to both practitioners and managers in data science.

The Art of Data Science: A Guide for Anyone Who Works with Data Details

Date   : Published September 5th 2015 by Leanpub
ISBN   :
Author : Roger D. Peng , Elizabeth Matsui
Format : ebook 154 pages
Genre   : Science, Nonfiction, Technology

[Download](#) The Art of Data Science: A Guide for Anyone Who Works w ...pdf

[Read Online](#) The Art of Data Science: A Guide for Anyone Who Works ...pdf

Download and Read Free Online The Art of Data Science: A Guide for Anyone Who Works with Data Roger D. Peng , Elizabeth Matsui

From Reader Review The Art of Data Science: A Guide for Anyone Who Works with Data for online ebook

Mustafa Hajmohammed says

Gentle introduction for those who are searching for more knowledge on data science in general and data mining in particular.
Not recommended for experienced data scientists, it is too basic.

Kristine says

This is an excellent book that breaks down the steps of data analysis. Many of these steps are carried out intuitively by data analysts, and it was enlightening to have them identified and put into context of the process. As a data analyst, I enjoyed this book. I think it would be a great read for people who work with analysts but don't have a clear view of what the work entails; I believe it would improve their appreciation of the creativity and complex processes carried out by the analysts with whom they collaborate.

Ayobami says

Great book for those new to data science. The illustrations and examples were apt.

Fermin Quant says

Great book with awesome examples. There are quite a few errors in grammar and spelling, but they do not subtract from the value of the knowledge. Very useful frameworks to understand and apply to data analyses. I did feel the last third part was a little rushed, and like it is somewhat incomplete, but overall the content was really good and useful.

Lolo says

I work with data, but I guess this book isn't for me.

I got the feeling that is was addressed to the experienced data scientist and not someone that wants to understand a little bit more about it. It seems that the author focuses more on the process and the logistics of the day-to-day tasks of a data scientist rather than the field of data science.

The book was interesting and well written but didn't really answer my questions. For sure I didn't learn the "art" of data science.

Javad Ebadi says

This book is very good to obtain a big picture about what is data analysis. The most important lesson that I learnt from this book was that a data analysis starts with a question, not with the data and at the end of the day it leads to another (better) question. The challenges in data analysis processes are described very well. Various types of question and types of data analysis are explained.

The book is very good for beginners, also it can be used by sophisticated data scientists.

Ahmad says

The authors share their experience in data analysis and the steps they propose seems necessary for a neat data analysis. I think I should re-read this book throughout my future data analysis projects.

Aina says

This excellent book takes you through each step of a typical data science project giving general advice, warning about common mistakes and giving many practical examples (including real industrial data science projects) to illustrate each of the points. It helps to build a mind map of options available at a data scientist's disposal during each of the project stages. Definitely a book to read multiple times.

Momen (????) says

It's a good book for anyone who wants to know more about data science and data science analysis
In this book, Roger D.peng showed the entire process of data science analysis :
1-stating a question
2-EDA (Exploratory data analysis)
3-Using Models & Expectations
4-inference and prediction
5- Interpreting Your Results
6-compunctions

Ayas Shaqour says

I wouldn't recommend this to someone who isn't an advanced or works as a data scientist.
Although I am pretty sure it would be too good for some one who Is/Does.

The book explains the untaught process, or lets say the unspoken process of a data scientist's job.
It explains the things you wouldn't read in a book of statistics, it explains the thought process taking place in a data scientist's mind.

It discusses the major steps taken to complete a task and how to judge every step you take.

It introduces the so called Epicycle and how it works.

But I will have to say this, If you don't work as a data scientist I think part of this book will confuse you or even scare you away. Some parts are easily understood through the first half of the book at least, but there are lots of unclear stuff and Jargons that will leave you really confused. Some examples where legit others again need way more advanced knowledge not because they are hard to understand but i felt they were a bit unclear or directed to let's say a "data scientist"

Again I am only saying that If you think this book is an introduction to data science in any kind then you might get a little disappointed.

Ray says

I read this in South Africa when it was recommended reading during the first week of a data science course. While it's hard to know for certain, I believe that the perspective I gained from it contributed towards my success in the course. I learned the most about all the non-science parts of a data scientist's job, and this was valuable to me later as I began to accumulate some of my earliest work experience in a data-science-like role.

Yahia El gamal says

This is a light data science book. I have mixed feelings towards it. Sometimes, Prof. Roger writes assuming some prior knowledge and experience (e.g. he wrote some parts freely assuming some knowledge about distributions, significance, ..) and those are the parts I really liked in the book. On the other hand, prof. Roger sometimes writes as if the audience are total newbies with no previous experience.

Leaving the above note aside, the abstraction Prof. Roger introduces about the process of data science / data analysis will probably be agreed upon by everyone. It probably won't add much to experienced people (as it is probably a second-nature) except when they start introducing data science and data analysis to others, and for this reason, i recommend this book. The abstractions and the examples will make introducing someone to the field a much organized process.

The book is highly accessible, enjoyable, and short. You won't invest much in reading it (time-wise) and you will get decent output from it.

Disclaimer: I read a pre-released version from leanpub

Ms_prue says

This book equipped me to answer all the questions I have in my data analyst life, specifically the "why am I here?" and "what is my purpose?" type ones. In my work, there are a lot of technical resources for data analysis tools but not a great deal of guidance on method. This book is exactly what I was looking for to fill that gap.

Kate says

This book was a mixed bag for me. Most of the time, I felt he was explaining basic research approaches (developing a hypothesis, exploratory data analysis, and especially the last chapter in presenting your findings). Other times I would get lost in his examples especially those without visualization. Overall, the book focused heavily on data analysis, and I think data science as a field also encompasses data extraction or collection, management, and virtualization. I would have liked to learn more about those topics.

Alex says

I'm somewhat ambivalent regarding this book -- I very much appreciate the pragmatic writing style, and there are some genuinely useful pieces of advice contained within. However, the target audience seems ambiguous. The best fit seems to be folks who are intending to take the JHU data science classes, and retrospectively this looks like it would be a very handy companion guide to the course. Having taken them, however, along with a number of other statistics/data analysis type classes, much of the content seems too cursory in review.