*P*REFACE

WHAT IS THE PRIMARY VALUE PROPOSITION FOR THIS BOOK?

This book contains a fast-paced introduction to relevant information about `Python`-based data visualization. You will learn how to generate graphics using `Pandas`, `Matplotlib`, and `Seaborn`. In addition, an appendix contains `SVG`-based and `D3`-based graphics effects, along with links for many additional code samples.

THE TARGET AUDIENCE

This book is intended primarily for those who have worked with `Python` and are interested in learning about graphics effects with `Python` libraries. It is also intended to reach an international audience of readers with highly diverse backgrounds in various age groups. Consequently, the book uses standard English rather than colloquial expressions that might be confusing to those readers. It provides a comfortable and meaningful learning experience for the intended readers.

WHAT WILL I LEARN?

The first chapter contains a quick tour of basic `Python` 3, followed by a chapter that introduces you to `NumPy`. The third and fourth chapters introduce you to `Pandas` as well as `Pandas` with `JSON` data. `MySQL` and `SQL`.

The fifth chapter delves into data visualization with `Matplotlib` and also working with SweetViz and Skimpy. The final chapter of this book shows you how to create graphics effects with `Seaborn`, and an example of a rendering graphics effects in `Bokeh`. In addition, an appendix is included with graphics effects based on `SVG` and `D3`.

WHY ARE THE CODE SAMPLES PRIMARILY IN PYTHON?

Most of the code samples are short (usually less than one page and sometimes less than half a page), and if need be, you can easily and quickly copy/paste the code into a

new Jupyter notebook. For the Python code samples that reference a CSV file, you do not need any additional code in the corresponding Jupyter notebook to access the CSV file. Moreover, the code samples execute quickly, so you won't need to avail yourself of the free GPU that is provided in Google Colaboratory.

If you do decide to use Google Colaboratory, you can easily copy/paste the Python code into a notebook, and also use the upload feature to upload existing Jupyter notebooks. Keep in mind the following point: if the Python code references a CSV file, make sure that you include the appropriate code snippet (details are available online) to access the CSV file in the corresponding Jupyter notebook in Google Colaboratory.

WHY DOES THIS BOOK INCLUDE SKLEARN MATERIAL?

First, keep in mind that the Sklearn material in this book is minimalistic, because it is not about machine learning. Second, the Sklearn material is located in chapter 6 where you will learn about some of the Sklearn built-in datasets. If you decide to study machine learning, you will have already been introduced to some aspects of Sklearn.

GETTING THE MOST FROM THIS BOOK

Some programmers learn well from prose, others learn well from sample code (and lots of it), which means that there's no single style that can be used for everyone.

Moreover, some programmers want to run the code first, see what it does, and then return to the code to delve into the details (and others use the opposite approach).

Consequently, there are various types of code samples in this book: some are short, some are long, and other code samples "build" from earlier code samples.

WHAT DO I NEED TO KNOW?

Current knowledge of Python 3.x is the most helpful skill. Knowledge of other programming languages (such as Java) can also be helpful because of the exposure to programming concepts and constructs. The less technical knowledge that you have, the more diligence will be required in order to understand the various topics that are covered.

As for the non-technical skills, it's very important to have a strong desire to learn about data visualization, along with the motivation and discipline to read and understand the code samples.

DON'T THE COMPANION FILES OBVIATE THE NEED FOR THIS BOOK?

The companion files contain all the code samples to save you time and effort from the error-prone process of manually typing code into a text file. In addition, there are situations in which you might not have easy access to these files. Furthermore, the code samples in the book provide explanations that are not available in the companion files.

DOES THIS BOOK CONTAIN PRODUCTION-LEVEL CODE SAMPLES?

The primary purpose of the code samples in this book is to show you `Python`-based libraries for data visualization. Clarity has higher priority than writing more compact code that is more difficult to understand (and possibly more prone to bugs). If you decide to use any of the code in this book in a production website, you ought to subject that code to the same rigorous analysis as the other parts of your code base.

HOW DO I SET UP A COMMAND SHELL?

If you are a Mac user, there are three ways to do so. The first method is to use `Finder` to navigate to `Applications > Utilities` and then double click on the `Utilities` application. Next, if you already have a command shell available, you can launch a new command shell by typing the following command:

```
open /Applications/Utilities/Terminal.app
```

A second method for Mac users is to open a new command shell on a Macbook from a command shell that is already visible simply by clicking `command+n` in that command shell, and your Mac will launch another command shell.

If you are a PC user, you can install Cygwin (open source *https://cygwin.com/*) that simulates bash commands, or use another toolkit such as MKS (a commercial product). Please read the online documentation that describes the download and installation process. Note that custom aliases are not automatically set if they are defined in a file other than the main start-up file (such as .bash_login).

COMPANION FILES

All the code samples and figures in this book may be obtained by writing to the publisher at *info@merclearning.com*.

WHAT ARE THE "NEXT STEPS" AFTER FINISHING THIS BOOK?

The answer to this question varies widely, mainly because the answer depends heavily on your objectives. If you are interested primarily in NLP, then you can learn more advanced concepts, such as attention, transformers, and the BERT-related models.

If you are primarily interested in machine learning, there are some subfields of machine learning, such as deep learning and reinforcement learning (and deep reinforcement learning) that might appeal to you. Fortunately, there are many resources available, and you can perform an Internet search for those resources. One other point: the aspects of machine learning for you to learn depend on who you are: the needs of a machine learning engineer, data scientist, manager, student or software developer are all different.

Oswald Campesato
September 2023