Preface

Over the years, I had to introduce a number of M.Sc. and Ph.D. students to the topic of process mining. Invariably, it was difficult to find a concise introduction to the topic, despite the fact that some of the fundamental ideas of process mining are quite simple. In principle, it should not be necessary to go through a series of research papers in order to get a good grasp of those ideas.

On the other hand, I did not want my students to start using ProM\(^1\) or Disco\(^2\) right away, without understanding what is happening behind the scenes. Instead, I would prefer to provide them with the working knowledge that would allow them to implement some simple process mining techniques on their own, even if only in a rudimentary form. It always seemed to me that being able to implement something is the best way to develop a solid understanding of a new topic.

The main goal of this book is to explain the core ideas of process mining and to show how these ideas can be implemented using just some basic tools that are available to any computer scientist or data scientist. One of such tools is the Python programming language, which has become very popular since it allows writing complex programs in a clear and concise form. Another tool that is very useful is the Graphviz library, which is able to display graphs and automatically calculate their layout without requiring the programmer to do so. Graphviz provides an effortless way to visualize the results of many process mining techniques.

Before going further, some disclaimers are in order; namely, this book is not meant to be a reference on process mining. In that sense, it would be very incomplete, since we will be using only a simplified version of a very small subset of process mining techniques. Also, the text does not delve into a wide variety of process models that can be generated by those techniques. Here, we will be using

\(^1\)http://www.promtools.org/.
\(^2\)https://fluxicon.com/disco/.
graphs (both directed and undirected, but just plain graphs) without getting into more sophisticated process modeling languages, such as Petri nets[^3] and BPMN[^4].

Nevertheless, this bare-bones approach should suffice to provide a feeling for what process mining is, while developing some skills that will definitely be useful in practice.

I prepared this text to be a very first introduction to process mining, and hence I called it a primer. After this, the reader can jump more confidently to the existing literature, namely the book by Wil van der Aalst[^5] and the extensive set of research publications in this field. I hope that this text will contribute towards a deeper understanding of process mining tools and techniques.

Lisbon, Portugal

Diogo R. Ferreira

February 2017

[^3]: http://www.informatik.uni-hamburg.de/TGI/PetriNets/
[^4]: http://www.bpmn.org/
[^5]: See [18] in the list of references on page 95.
A Primer on Process Mining
Practical Skills with Python and Graphviz
Ferreira, D.R.
2017, VIII, 96 p. 37 illus., 15 illus. in color., Softcover
ISBN: 978-3-319-56426-5