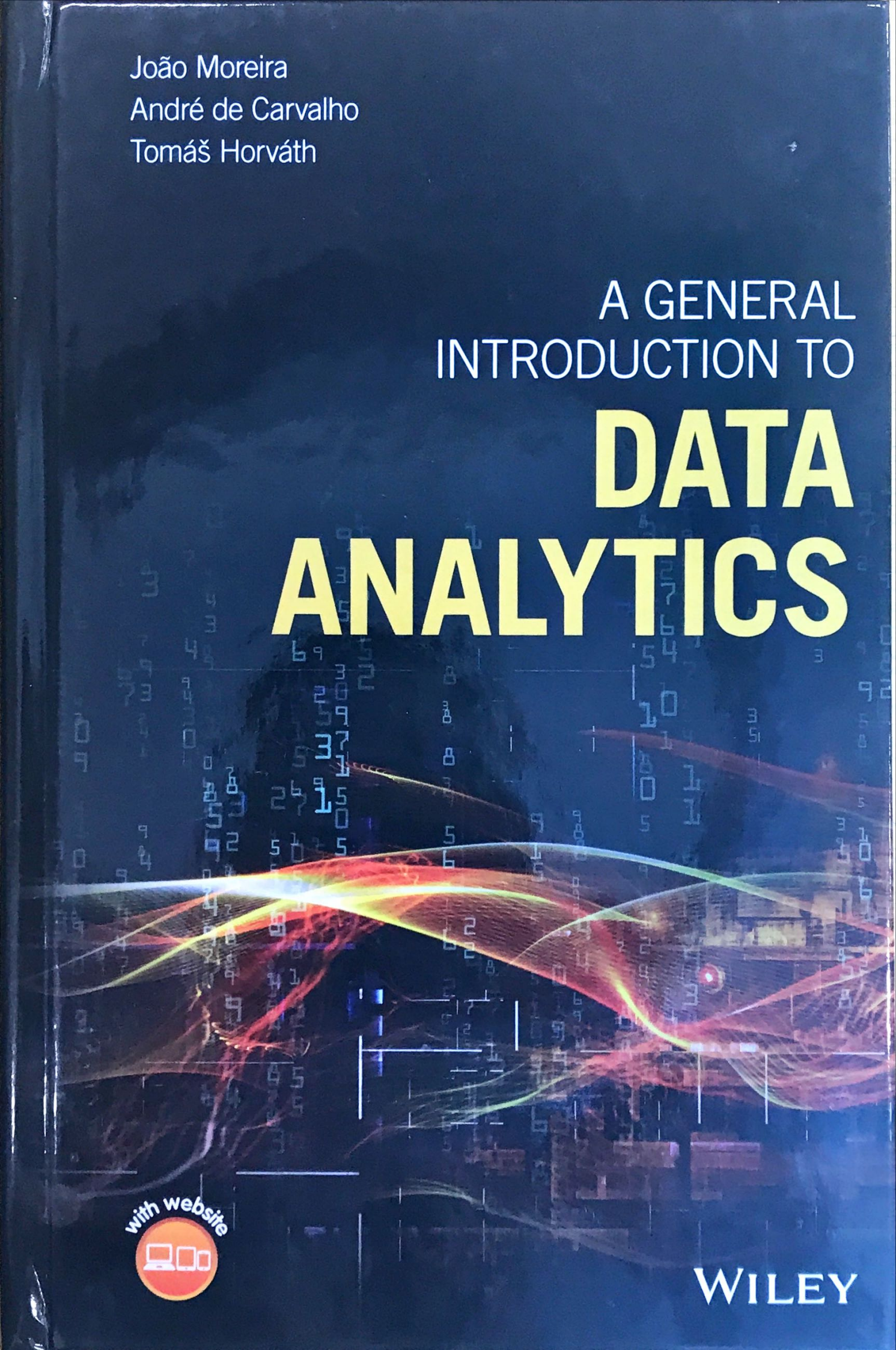
João Moreira
André de Carvalho
Tomáš Horváth

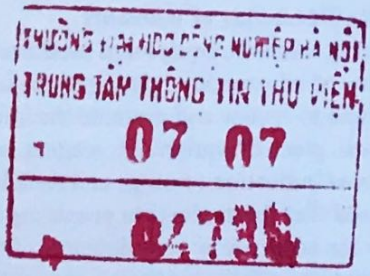
A GENERAL INTRODUCTION TO DATA ANALYTICS



WILEY

A General Introduction to Data Analytics

João Mendes Moreira
University of Porto



André C. P. L. F. de Carvalho
University of São Paulo

Tomáš Horváth
Eötvös Loránd University in Budapest
Pavol Jozef Šafárik University in Košice

WILEY

This edition first published 2019

Registered Office
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office
111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Library of Congress Cataloging-in-Publication Data

Names: Moreira, João, 1969– author. | Carvalho, André Carlos Ponce de Leon Ferreira, author. | Horváth, Tomáš, 1976– author.
Title: A general introduction to data analytics / by João Mendes Moreira, André C. P. L. F. de Carvalho, Tomáš Horváth.
Description: Hoboken, NJ : John Wiley & Sons, 2019. | Includes bibliographical references and index. |
Identifiers: LCCN 2017060728 (print) | LCCN 2018005929 (ebook) | ISBN 9781119296256 (pdf) | ISBN 9781119296263 (epub) | ISBN 9781119296249 (cloth)
Subjects: LCSH: Mathematical statistics–Methodology. | Electronic data processing. | Data mining.
Classification: LCC QA276.4 (ebook) | LCC QA276.4 .M664 2018 (print) | DDC 519.50285–dc23
LC record available at https://lccn.loc.gov/2017060728

Cover image: © agsandrew/Shutterstock
Cover design by Wiley

Printed in the United States of America.

Set in 10/12pt Warnock by SPi Global, Pondicherry, India

V10012656_073019

Contents

A guide to the principles and methods of data analysis that does not require knowledge of statistics or programming

A General Introduction to Data Analytics is an essential guide to understand and use data analytics. This book is written using easy-to-understand terms and does not require familiarity with statistics or programming. The authors—noted experts in the field—highlight an explanation of the intuition behind the basic data analytics techniques. The text also contains exercises and illustrative examples.

Thought to be easily accessible to non-experts, the book provides motivation to the necessity of analyzing data. It explains how to visualize and summarize data, and how to find natural groups and frequent patterns in a dataset. The book also explores predictive tasks, be them classification or regression. Finally, the book discusses popular data analytic applications, like mining the web, information retrieval, social network analysis, working with text, and recommender systems. The learning resources offer:

- A guide to the reasoning behind data mining techniques
- A unique illustrative example that extends throughout all the chapters
- Exercises at the end of each chapter and larger projects at the end of parts II and III of the book
- Supplemented with PowerPoint slides available for instructors on a Wiley Book Companion Site

Together with these learning resources, the book can be used in a 13-week course guide, one chapter per course topic.

The book was written in a format that allows the understanding of the main data analytics concepts by non-mathematicians, non-statisticians and non-computer scientists interested in getting an introduction to data science. *A General Introduction to Data Analytics* is a basic guide to data analytics written in highly accessible terms.

João Mendes Moreira, PhD, is an assistant professor in the Faculty of Engineering at the University of Porto, Porto, Portugal and is also a researcher in LIAAD-INESC TEC, Porto, Portugal.

André de Carvalho, PhD, is a full professor in the Institute of Mathematics and Computer Science at the University of São Paulo, Brazil.

Tomáš Horváth, PhD, is an assistant professor at the Faculty of Informatics of the Eötvös Loránd University in Budapest, Hungary, and is also associated with the Faculty of Science at the Pavol Jozef Šafárik University in Košice, Slovakia.

A companion website with additional resources is available at:
www.wiley.com/go/moreira/dataanalytics

ISBN 978-1-119-29624-9

Cover Design: Wiley
Cover Image: © agsandrew/Shutterstock

WILEY

Also available as an e-book