International Series in Operations Research & Management Science

Volume 264

Series Editor

Camille C. Price
Stephen F. Austin State University, TX, USA

Associate Series Editor

Joe Zhu
Worcester Polytechnic Institute, MA, USA

Founding Series Editor

Frederick S. Hillier, Stanford University, CA, USA

More information about this series at http://www.springer.com/series/6161

Bhimasankaram Pochiraju • Sridhar Seshadri
Editors

Essentials of Business Analytics

An Introduction to the Methodology and its Applications

Springer

Editors
Bhimasankaram Pochiraju
Applied Statistics and Computing Lab
Indian School of Business
Hyderabad, Telangana, India

Sridhar Seshadri
Gies College of Business
University of Illinois at Urbana Champaign
Champaign, IL, USA

ISSN 0884-8289 ISSN 2214-7934 (electronic)
International Series in Operations Research & Management Science
ISBN 978-3-319-68836-7 ISBN 978-3-319-68837-4 (eBook)
https://doi.org/10.1007/978-3-319-68837-4

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Professor Bhimasankaram: With the divine blessings of Bhagawan Sri Sri Sri Satya Sai Baba, I dedicate this book to my parents—Sri Pochiraju Rama Rao and Smt. Venkata Ratnamma.

Sridhar Seshadri: I dedicate this book to the memory of my parents, Smt. Ranganayaki and Sri Desikachari Seshadri, my father-in-law, Sri Kalyana Srinivasan Ayodhyanath, and my dear friend, collaborator and advisor, Professor Bhimasankaram.

Contents

1 Introduction 1
Sridhar Seshadri

Part I Tools

2 Data Collection 19
Sudhir Voleti

3 Data Management—Relational Database Systems (RDBMS) 41
Hemanth Kumar Dasararaju and Peeyush Taori

4 Big Data Management 71
Peeyush Taori and Hemanth Kumar Dasararaju

5 Data Visualization 111
John F. Tripp

6 Statistical Methods: Basic Inferences 137
Vishnuprasad Nagadevara

7 Statistical Methods: Regression Analysis 179
Bhimasankaram Pochiraju and Hema Sri Sai Kollipara

8 Advanced Regression Analysis 247
Vishnuprasad Nagadevara

9 Text Analytics 283
Sudhir Voleti

Part II Modeling Methods

10 Simulation 305
Sumit Kunnumkal

11 Introduction to Optimization 337
Milind G. Sohoni

12 Forecasting Analytics 381
Konstantinos I. Nikolopoulos and Dimitrios D. Thomakos

13 Count Data Regression 421
Thriyambakam Krishnan

14 Survival Analysis 439
Thriyambakam Krishnan

15 Machine Learning (Unsupervised) 459
Shailesh Kumar

16 Machine Learning (Supervised) 507
Shailesh Kumar

17 Deep Learning 569
Manish Gupta

Part III Applications

18 Retail Analytics 599
Ramandeep S. Randhawa

19 Marketing Analytics 623
S. Arunachalam and Amalesh Sharma

20 Financial Analytics 659
Krishnamurthy Vaidyanathan

21 Social Media and Web Analytics 719
Vishnuprasad Nagadevara

22 Healthcare Analytics 765
Maqbool (Mac) Dada and Chester Chambers

23 Pricing Analytics 793
Kalyan Talluri and Sridhar Seshadri

24 Supply Chain Analytics 823
Yao Zhao

25 Case Study: Ideal Insurance 847
Deepak Agrawal and Soumithri Mamidipudi

26 Case Study: AAA Airline 863
Deepak Agrawal, Hema Sri Sai Kollipara, and Soumithri Mamidipudi

27 Case Study: InfoMedia Solutions 873
Deepak Agrawal, Soumithri Mamidipudi, and Sriram Padmanabhan

28 Introduction to R 889
Peeyush Taori and Hemanth Kumar Dasararaju

29 Introduction to Python 917
Peeyush Taori and Hemanth Kumar Dasararaju

30 Probability and Statistics 945
Peeyush Taori, Soumithri Mamidipudi, and Deepak Agrawal

Index 965

Disclaimer

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Acknowledgements

This book is the outcome of a truly collaborative effort amongst many people who have contributed in different ways. We are deeply thankful to all the contributing authors for their ideas and support. The book belongs to them. This book would not have been possible without the help of *Deepak Agrawal*. *Deepak* helped in every way, from editorial work, solution support, programming help, to coordination with authors and researchers, and many more things. *Soumithri Mamidipudi* provided editorial support, helped with writing summaries of every chapter, and proof-edited the probability and statistics appendix and cases. *Padmavati Sridhar* provided editorial support for many chapters. Two associate alumni—*Ramakrishna Vempati* and *Suryanarayana Ambatipudi*—of the *Certificate Programme in Business Analytics* (CBA) at *Indian School of Business* (ISB) helped with locating contemporary examples and references. They suggested examples for the Retail Analytics and Supply Chain Analytics chapters. *Ramakrishna* also contributed to the draft of the Big Data chapter. Several researchers in the *Advanced Statistics and Computing Lab* (ASC Lab) at ISB helped in many ways. *Hema Sri Sai Kollipara* provided support for the cases, exercises, and technical and statistics support for various chapters. *Aditya Taori* helped with examples for the machine learning chapters and exercises. *Saurabh Jugalkishor* contributed examples for the machine learning chapters. The ASC Lab's researchers and *Hemanth Kumar* provided technical support in preparing solutions for various examples referred in the chapters. *Ashish Khandelwal, Fellow Program student at ISB*, helped with the chapter on Linear Regression. *Dr. Kumar Eswaran* and *Joy Mustafi* provided additional thoughts for the Unsupervised Learning chapter. The editorial team comprising *Faith Su*, *Mathew Amboy* and series editor *Camille Price* gave immense support during the book proposal stage, guidance during editing, production, etc. The ASC Lab provided the research support for this project.

We thank our families for the constant support during the 2-year long project. We thank each and every person associated with us during the beautiful journey of writing this book.

Contributors

Deepak Agrawal Indian School of Business, Hyderabad, Telangana, India

S. Arunachalam Indian School of Business, Hyderabad, Telangana, India

Chester Chambers Carey Business School, Johns Hopkins University, Baltimore, MD, USA

Maqbool (Mac) Dada Carey Business School, Johns Hopkins University, Baltimore, MD, USA

Manish Gupta Microsoft Corporation, Hyderabad, India

Hema Sri Sai Kollipara Indian School of Business, Hyderabad, Telangana, India

Thriyambakam Krishnan Chennai Mathematical Institute, Chennai, India

Shailesh Kumar Reliance Jio, Navi Mumbai, Maharashtra, India

Hemanth Kumar Dasararaju Indian School of Business, Hyderabad, Telangana, India

Sumit Kunnumkal Indian School of Business, Hyderabad, Telangana, India

Soumithri Mamidipudi Indian School of Business, Hyderabad, Telangana, India

Vishnuprasad Nagadevara IIM-Bangalore, Bengaluru, Karnataka, India

Konstantinos I. Nikolopoulos Bangor Business School, Bangor, Gwynedd, UK

Sriram Padmanabhan New York, NY, USA

Bhimasankaram Pochiraju Applied Statistics and Computing Lab, Indian School of Business, Hyderabad, Telangana, India

Ramandeep S. Randhawa Marshall School of Business, University of Southern California, Los Angeles, CA, USA

Sridhar Seshadri Gies College of Business, University of Illinois at Urbana Champaign, Champaign, IL, USA

Amalesh Sharma Texas A&M University, College Station, TX, USA

Milind G. Sohoni Indian School of Business, Hyderabad, Telangana, India

Kalyan Talluri Imperial College Business School, South Kensington, London, UK

Peeyush Taori London Business School, London, UK

Dimitrios D. Thomakos University of Peloponnese, Tripoli, Greece

John F. Tripp Clemson University, Clemson, SC, USA

Krishnamurthy Vaidyanathan Indian School of Business, Hyderabad, Telangana, India

Sudhir Voleti Indian School of Business, Hyderabad, Telangana, India

Yao Zhao Rutgers University, Newark, NJ, USA