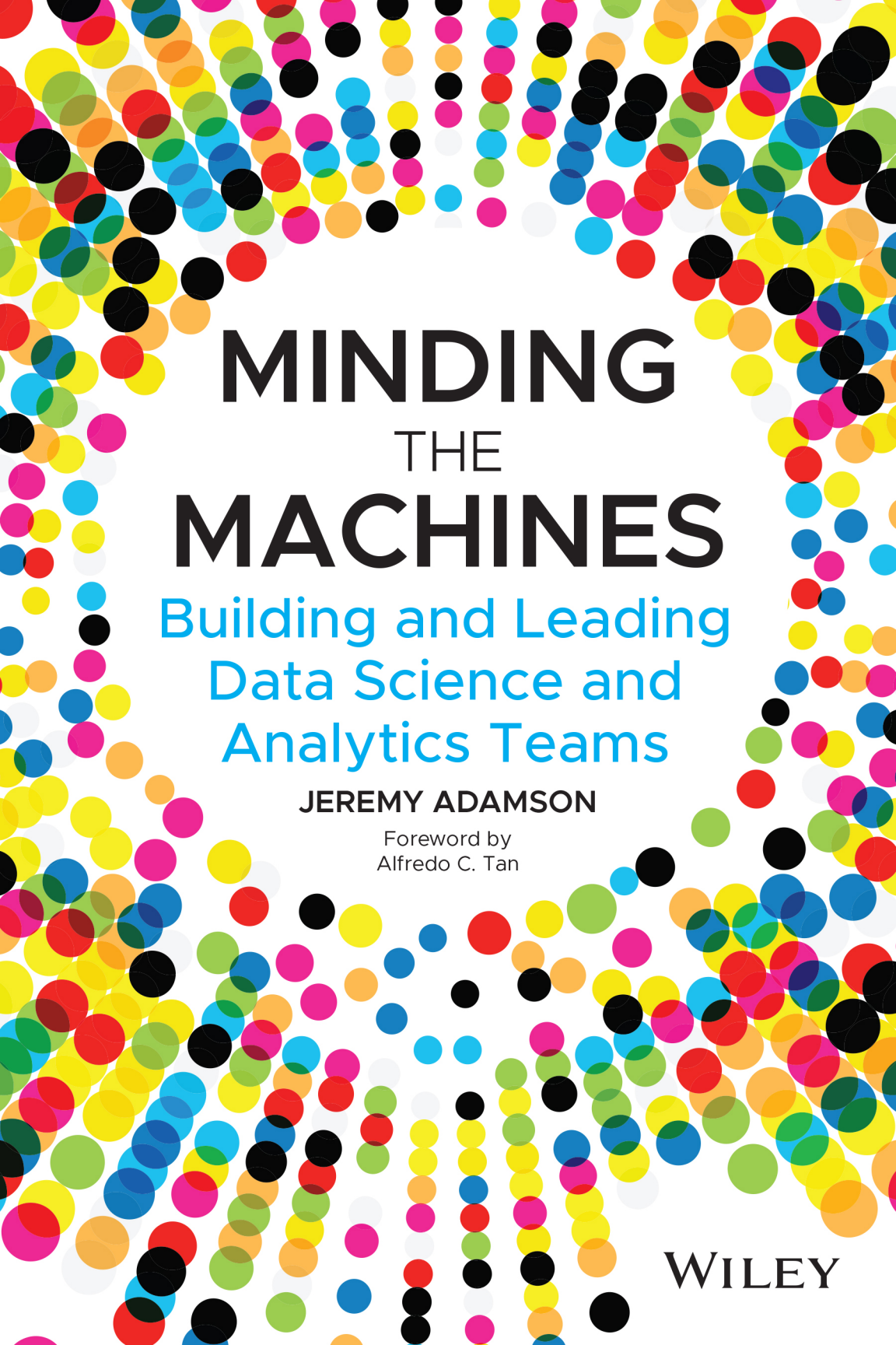
MINDING
THE
MACHINES
Building and Leading
Data Science and
Analytics Teams
JEREMY ADAMSON
Foreword by
Alfredo C. Tan
WILEY

Minding the Machines

Building and Leading Data Science and Analytics Teams

Jeremy Adamson

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

ISBN: 978-1-119-78532-3

ISBN: 978-1-119-78534-7 (ebk)

ISBN: 978-1-119-78533-0 (ebk)

Library of Congress Control Number: 2021938292

Cover Image: © shuoshu/Getty Images
Cover Design: Wiley

to Melo

About the Author

Jeremy Adamson is a leader in AI and analytics strategy and has a broad range of experience in aviation, energy, financial services, and public administration. Jeremy has worked with several major organizations to help them establish a leadership position in data science and to unlock real business value using advanced analytics.

Jeremy holds a master's degree in transportation engineering and a bachelor's degree in civil engineering from the University of New Brunswick as well as a master of business administration from the University of Calgary. Jeremy is a professional engineer in the province of Alberta. He currently resides in Calgary.

Learn more at `www.rjeremyadamson.com`.

About the Technical Editor

Andrew McHardy is a strategic leader and advisor with over 25 years' experience in analytics and artificial intelligence, engineering, and management consulting. He has held roles ranging from field to executive level, working with asset-intensive organizations in Canada and internationally, across the energy, resources, manufacturing, transportation, and defense sectors. Throughout his career, he has focused on leveraging digital tools, predictive analytics, and artificial intelligence to drive improvements in workforce safety and productivity, capital project delivery, asset reliability and maintenance, quality, and production.

Andrew holds a master's degree in Statistics and Industrial and Systems Engineering from Rutgers University and a bachelor's degree in Chemical and Materials Engineering and master's degree in Defence Studies from the Royal Military College of Canada. He has also completed a business analytics specialization through the University of Pennsylvania, and is a certified Project Management Professional.

Prior to transitioning to the private sector, Andrew served in the Canadian Army and Special Operations Forces and has deployed on operations in Canada and internationally.

About the Foreword Author

A thought leader, change agent, and impact-focused innovator, Alfredo C. Tan is an industry professor in the Strategic Management faculty at McMaster University where he teaches at the undergraduate and graduate level and in The Directors College Innovation Governance program for board directors. He is also the senior vice president of Strategy, Data and Products at Rogers Sports & Media, where his team is helping lead the digital transformation of the organization.

Alfredo has led companies through changing digital landscapes and evolving consumer expectations and behaviors. Most recently he was the chief marketing officer and the first Chief Digital and Innovation Officer at WestJet Airlines. Alfredo is also a former Facebook and Instagram and Yahoo! executive, having worked in Canada, the United States, Asia-Pacific, and Latin America. Mr. Tan is a frequently requested keynote speaker at leading technology, media, marketing, corporate, and various other industry conferences. He has been profiled and quoted in *Forbes Magazine*, *Marketing Magazine*, the *Economist*, *Strategy Magazine*, the *Globe and Mail*, and many other leading industry publications. Alfredo currently sits on the Advisory Board on Innovation at McMaster University, is a member of the Leadership Council at the Perimeter Institute for Theoretical Physics, and sits on the Haskayne School of Business Technology Management Industry Advisory Council at the University of Calgary. Alfredo was recognized with the Desmond Parker Outstanding Young Alumni Award from the University of Toronto, has been inducted into the McMaster University Alumni Gallery, and is a recipient of the prestigious

Arbor Award for his service to the University of Toronto. Alfredo holds an HBSc in Forensic Science & Biology from the University of Toronto, an MBA from McMaster University, and an LLM from the University of Toronto Law School. He is a graduate of the Advanced Management Program at the Harvard Business School.

He and his partner, Jessica, are proud parents to five rescue animals and hope to own a large property one day where all their rescue animals can roam free.

Acknowledgments

I have been fortunate to work with many amazing people and organizations in the field of analytics. It was alongside them that the lessons of this book were learned. It would be impossible to list them all and their contributions, but to all my past and present colleagues, my sincerest appreciation for the support throughout the writing of this book, for your valuable feedback, and for your patience as we struggled together to figure out how best to build out a strong and value-driven analytics function.

Thank you to Devon Lewis, my acquisitions editor, who believed in this project and helped me to navigate the process of writing a book.

Thank you to Christine O'Connor, the project manager, and to Judy Flynn and Janet Wehner, the editors, for helping to wrangle my writing and keep me on track.

Thank you to Andrew McHardy for agreeing to be the technical editor for this project. His wisdom and experience really elevated the material and helped to focus the book on core concepts. He is the embodiment of parsimony.

Thank you to Alfredo Tan for his ongoing support and for contributing a foreword and sharing his experience in the field.

Finally, thank you to my wife, Melodie, and my boys, Finley and Wyatt, for all their patience and support through this project. My appreciation in particular to Finley and Wyatt for their careful review and approval of the swimlane and Venn diagrams that appear in the book.

—Jeremy Adamson

Contents at a Glance

Contents

Foreword

Data. There was a time when this word made reference to a *Star Trek* character or something professionals in the IT department who worked on databases would manage. Today *data*, *data science*, *data engineering*, *data analysts*, or any term including the use of *data* is pervasive across business, industries, and society. The use of the term *data* has practically become everyday vernacular in business; it seems to be the holy grail solution to everything. However, most organizations are still in the very early stages of their journey.

Many of the world's leading organizations can attribute their success to the fact that the practice of data science is increasingly becoming a strategic function. Analytics and data science enable consumer experiences that have become indispensable in our daily lives and deliver highly personalized recommendations and content, and this is now the expectation for almost everything else in our lives. The expectation of the customer has become immediate, personalized services that predict what it is they may want before they may even know it themselves. Data is what powers these great product experiences. Data science is no longer simply a technology function buried within IT or reserved purely for the tech giants in Silicon Valley. Data science and analytics will become increasingly indispensable in health care as it will improve diagnostic accuracy and efficiency. In finance, it will aid in the detection of anomalies and fraud. In manufacturing, it will aid in fault prediction and preventative maintenance. Whether you work in corporate strategy, research & insights, product development, human resources, marketing,

technology, or finance, you will no longer be able to effectively compete without leveraging the talent and capabilities of the data science teams.

The need for knowledge in Data Science & Analytics, Algorithms & Artificial Intelligence is becoming evident in the sheer volume of online courses, degrees, and certifications available on EDx, Coursera, Udacity, and other online education providers. Top-ranked universities across Canada have introduced graduate degrees in data science and analytics. Two of the most prestigious universities in the world, the University of California, Berkeley, and Massachusetts Institute of Technology, are creating entirely new institutions within their campuses to come to terms with the ubiquity of data and the rise of artificial intelligence.

However, it isn't simply technical, mathematical, or scientific horsepower that is required by organizations in the data science world. In most organizations the premise is still that data science teams are over-indexed in the technical practice versus being embedded in the business to drive business performance. The most successful data science teams are those that have a focus on contributing to the strategy of hiring and retaining people who are focused on value creation and finding ways to democratize access to data and decision making. Because it is one of the newest functions in most organizations, there is little body of work to refer to on how to design and build the right data science organization. We are all learning in real time, across all industries and geographies. How do you hire? How do you structure the teams? What problems do you solve? How do you set up the culture of experimentation? How do you think about democratizing access? How do you evolve beyond reporting and move into prediction models and algorithms?

I have had the pleasure of being a senior leader to technology teams at organizations with widely varying analytical maturity, in industries ranging from Silicon Valley giants to aviation and from sports to media and digital. The key differentiator for those teams finding success using data science is not whether they have a data lake or are deploying neural nets and reinforcement learning. The winning teams are those that integrate into the organization, understand the business, build strong relationships, collaborate, align their objectives to the business, and see data science as a toolkit for solving business problems rather than an esoteric and technical field of study. They are integrated into the business and serve a strategic function with support at the highest levels of the organization, all the way to the president or CEO. The three pillars described in this book, people, process, and strategy, are every bit as important as the data and technology. The challenge, despite all the focus on the technical skills, is still a very human one. There is no

doubt machines are helping drive more automation and the increasing power of data and algorithms to help make decisions. I would argue that the importance of the human element, the people responsible for building the models, doing the analysis and creating the algorithms, will becoming increasingly important. The need for leadership, empathy, and an understanding of organizational behavior is becoming increasingly important. It is essential for these teams to have the ability to deal across the enterprise with privacy and data governance, ethics, and bias, and to ensure that the capital and operational investments are solving the problems that really matter. As the field of data science advances, the human element and the team you build becomes even more important. The more important the machines become, and the data that powers them, the more the people element will be critical. Strategy, process, culture, and the human side of data science will be the next evolution of the practice to deliver on the promise of big data and business results.

Most data science leaders, focusing mainly on the technical aspects of their craft, have struggled to find successes in organizations and to unlock real business value. *Minding the Machines* helps to fill that gap and redirect these professionals to the things that matter. Blending the science of data and the leadership of people, process, and strategy is what Jeremy manages to do brilliantly in this book.

—Alfredo C. Tan

Introduction

Minding the Machines provides insights into how to structure and lead a successful analytics practice. Establishing this practice requires a significant up-front investment in understanding and contextualizing the initiative in contrast to better-understood functions such as IT or HR. Many organizations have attempted to use operating models and templates from these other functions, showing a fundamental misunderstanding of where analytics fits within an organization and leading to visible failures. These failures have set back the analytical maturity of many organizations. Business leaders need to hire or develop data-centric talent who can step back from analysis and project management to view their work through a lens of value creation.

Readers will understand how organizations and practitioners need to structure, build, and lead a successful analytics team—to bridge the gap between business leaders and the analytical function. The analytics job market is booming, and the talent pool has swelled with other professionals upskilling and rebranding themselves as data scientists. While this influx of highly technical specialists with limited leadership experience has had negative consequences for the practice, it also provides an opportunity for personal differentiation.

Minding the Machines is organized in three key pillars: strategy, process, and people.

Strategy—How to assess organizational readiness, identify gaps, establish an attainable roadmap, and properly articulate a value proposition and case for change.

Process—How to select and manage projects across their life cycle, including design thinking, risk assessment, governance, and operationalization.

People—How to structure and engage a team, establish productive and parsimonious conventions, and lead a distinct practice with unique requirements.

Minding the Machines is intended for analytics practitioners seeking career progression, business leaders who wish to understand how to manage this unique practice, and students who want to differentiate themselves against their technical peers.

There is a significant need for leaders who can bridge the gap between the business and the data science and analytics functions. *Minding the Machines* fills this need, helping data science professionals to successfully leverage this powerful practice to unlock value in their organizations.

How to Contact the Publisher

If you believe you've found a mistake in this book, please bring it to our attention. At John Wiley & Sons, we understand how important it is to provide our customers with accurate content, but even with our best efforts an error may occur.

In order to submit your possible errata, please email it to our Customer Service Team at `wileysupport@wiley.com` with the subject line "Possible Book Errata Submission."

How to Contact the Author

I would love to connect and hear what you thought of this book or to discuss opportunities to collaborate. You can reach me via:

Website: `www.rjeremyadamson.com`
Email: `jeremy@rjeremyadamson.com`
LinkedIn: `https://linkedin.com/in/rjeremyadamson/`
Twitter: `@r2b7e`
Instagram: `r2b7e`

CHAPTER

1

Prologue

How is analytics unique in a corporate context? What have other organizations done right? What have they done wrong? What are the expectations for a new analytics leader?

Building, integrating, and leading effective analytics teams is a business imperative. The organizations that are most successful overall are those that effectively leverage their analytics capabilities to build a sustainable competitive advantage. However, many organizations are simply not getting the return that they expected on their investments in analytics.

Does hiring an engineer cause the surrounding buildings to be more robust? Could hiring five engineers make those buildings even more robust? Would hiring a pharmacist make you healthier? Would hiring an actuary increase your longevity?

In the last 20 years, the once sleepy academic fields of statistics, operations research, and decision support have exploded and been rebranded using terms such as *data science*, *artificial intelligence*, *big data*, and *advanced analytics*, among others. These practices have matured, and they have entered the mainstream. Advanced analytics and AI have moved from

being an investment in the future to a core component of corporate strategy and a key enabler for all areas of the business. Regardless of their size or the industry or sector in which they are used, these technologies are becoming an organizational necessity throughout the world. Organizations use advanced analytics and AI to develop new products, understand their customers, control their costs, and make better-informed decisions.

This new corporate function has been integrated in several support and core functions and has quickly become indispensable. Analytics is expected to add $16 trillion US to the global economy by 2030 and companies are eager to realize some of that value (PwC, 2017). As a result there has been a surge in demand for practitioners. There are approximately 3.3 million people employed in analytics in North America, and this is projected to grow by 15 percent a year in the United States over the next decade according to the US Bureau of Labor Statistics (2020). Educational institutions are eager to meet this demand.

Essentially every major college and university in the world offers some sort of analytics program or specialization, within multiple faculties such as mathematics, engineering, or business. There are several hundred books published in this space, and it enjoys a highly active online community. These resources are strong, edifying, and comprehensive and cover every new technology, framework, algorithm, and approach. With such an active community, new algorithms and methodologies are packaged and made publicly accessible for tools such as R, Python, and Julia, almost immediately after being developed. The best and brightest are choosing to enter the field, often called the "Sexiest Job of the 21st Century" (Davenport & Patil, 2012).

So, with overwhelming demand and a staggeringly capable pool of talent, why are there so many failures? Why are most organizations struggling to unlock the value in data science and advanced analytics? With so much executive support, so much talent, so much academic focus, why are so few organizations successfully deploying and leveraging analytics? In the 1980s, economist Robert Solow remarked that "you can see the computer age everywhere except in the productivity statistics." Why now can we see data science transforming organizations without a commensurate improvement in productivity?

The fundamental reason is in the opening questions. Clearly, having five engineers on your payroll will not improve buildings in the vicinity. Even if directly requested, ordered, mandated, or incented to "improve a building," they will struggle to do so. Replaced with any other profession, it is clear why this cannot work, but it persists as the typical furtive first steps into analytics. This is the essence of what most organizations

have done with their data analytics team. They have hired talented, passionate, ambitious data scientists and asked them to simply "do data science." Advanced analytics and AI as a practice has been poorly defined in its scope, has been subjected to great overspending, and has been lacking in relevant performance measures. Because of this, it is failing to live up to its potential.

Effectively all organizations realize the benefits of analytics. In a survey by Deloitte in 2020, 43 percent believed their organization would be transformed by analytics within the next 1 to 3 years, and 23 percent within the next year (Ammanath, Jarvis, & Hupfer, 2020). Though most organizations are on board with analytics being a key strategic advantage, they are unaware of how exactly to extract value from the new function.

Short-tenured data scientists, employed in a frothy and competitive market, share stories of unfocused and baffled companies where they have been engaged in operational reporting, confirming executive assumptions, and adding visualizations to legacy reports. Uncertain what to do with the team, and in a final act of surrender, the companies no longer expect the function to "do data science" and transform the team into a disbanded group of de facto technical resources automating onerous spreadsheets in a quasi-IT role.

Contrasted with those organizations who have truly got it right, the differences are stunning. For several companies, well-supported analytical Centers of Excellence are a key team, perpetually hiring and growing, and are solicited for their input and perspectives on all major projects. In others, internal Communities of Practice encourage cross-pollination of ideas and development opportunities for junior data scientists. New products are formed and informed after a thorough analysis by data scientists, who are also supporting human resources with success indicators and spending Fridays pursuing their transformative passion projects. Theories and hypotheses are quickly tested in a cross-functional analytical sandbox. Individuals are sharing their work at conferences and symposia, building eminence, and gaining acclaim for the organization.

The irony is that while most analytics teams exercise great care in deconstructing a problem, modeling each of the elements, and developing robust simulations and marvelously elegant solutions, the business processes that they employ in their project delivery are often at best immature and at worst destructive. The challenge for any leader is to encourage and stimulate logical thinking from a lens of interconnectedness and value creation, and to direct that philosophy to execution as well. As the logician Bertrand Russell said, "What is best in mathematics deserves not merely to be learnt as a task, but to be assimilated as a part

of daily thought, and brought again and again before the mind with ever-renewed encouragement" (Russell, 1902).

The title of this book, *Minding the Machines*, is meant to be an affectionate recursion. Those talented and creative practitioners, craftspeople, data scientists, and machine learning engineers, who create the algorithms that are transforming the way business is done, mind and care for those machines like a shepherd. Those machines need to be trained, informed, given established processes, encouraged to be broadly interoperable, and developed to be applicable to many different situations and problems. Similarly, the practitioners themselves need to be minded, cared for, cultivated, and encouraged for both the team and the individuals to be successful. This concept of recursion occurs throughout the book.

The second key theme is one of parsimony. Parsimony is a philosophy of intentionally expending the minimum amount of energy required in an activity so as to maintain overall efficiency. This is a key part of modeling; it is about keeping things as simple as possible, but no simpler. Similarly, and in the vein of recursion, teams themselves must be parsimonious. For analytics to mature as a practice while still delivering an accelerated time-to-value, teams need to be scrappy and lean.

At a foundational level, the objective of this book is to provide clear insights into how to structure and lead a successful analytics team. This is a deceptively challenging objective since there are no generalized templates from which to work. Establishing a project management office, information services, or human resources department is an understood process and does not vary materially between organizations. Establishing an analytics team, by contrast, requires a significant up-front investment in understanding and contextualizing the initiative. Many organizations have attempted to use operating models and templates from other functions—often IT and operations research. This fundamental misunderstanding of where analytics fits within an organization has led to visible failures and has set back the analytical maturity of many organizations. Business leaders need to hire or develop value-centric talent who can step back from analysis and project management to view their work as existing within a network of individuals and teams with competing priorities and motivations.

Corporations, without a template or a default methodology to benchmark against, have made expensive missteps in building these teams by installing the wrong leaders, copying other functions, positioning the function under the wrong executives, incentivizing destructive behaviors, hiring the wrong people, and committing to the wrong projects. They have contracted expensive consultants to provide roadmaps that do not consider the unique culture or competencies of their organization.

They have embedded disparate data science specialists throughout the organization and incurred enormous technical debt.

These issues are not insurmountable, however. Whether an organization is beginning its first foray into the analytics space or it is rebooting a failed team for the third time, the key is the creation of a carefully considered strategy, the establishment of realistic goals, and the full commitment of executive leadership. The advantages associated with analytics are too great to overlook, and the long-term cumulative impact of interrelated and interdependent models provides a powerful incentive for aggressive adoption.

This book was written for anybody who aspires to lead or be part of an effective analytics team, regardless of managerial experience. Every analytics leader, from a first-time team lead to a seasoned VP, has unique challenges to overcome.

For the Leader from the Business

Every new role is a challenge, regardless of ability, disposition, or motivation. This is particularly the case with a unique subculture of academic technocrats with whom it is difficult to establish credibility without enough time being "hands on keyboard." Without the respect of your team, it is impossible to get the buy-in required to establish best practices and ensure that the output of the team is not simply self-satisfying experimentation but can bring real value to the organization. As a corollary, every practitioner has experienced a manager who is out of their depth and who has compensated for their lack of self-confidence with authoritarianism and distrust, shifting the focus of the team toward an end they are more confident with, such as reporting.

For all but the most analytically committed organizations, there is a point along the chain of command where a practitioner reports to a non-practitioner. This can be a challenging junction for both parties without clear expectations, transparent communication, and mutual respect. Catching up from a technical perspective isn't feasible or advisable, but by leveraging your business understanding and domain knowledge to become an intermediary, translating business needs into projects and analytical outputs into operationalizable processes, you can unlock the power of your new team and give them the opportunity to develop into more business-oriented individuals.

For the business leader, I hope that this book helps you to reframe and refine your current leadership abilities, and to use them in an analytics context in order to engage your team and find success together.