Introduction

Being a data scientist,[2] as it turns out, is a lot more complicated than you and I had realized when we were taking courses in machine learning (ML) and programming. Once we started working, we found it was not actually the algorithms or the coding that was so complicated.

Instead, it was the business aspects of our jobs—the nontechnical competencies and organizational savvy that we didn't cover in our coursework—that turned out to be difficult for us. And for data scientists, business skills are the most critical to success in non-academic settings.

In my own career, as I have managed, trained and worked alongside hundreds of data scientists over the past two decades, I don't remember ever hearing a complaint that the algorithms were too difficult. However, I've often listened as data scientists described struggles within the workplace—getting burned by office politics, being assigned mind-numbingly boring tasks, aiming at moving targets, or working with unclear expectations from leadership.

On the flip side, I've also listened to business executives and senior managers moan, in private, about how their data scientists are not effectively communicating, managing projects, relating to stake-holders, or building trust. They've lamented the difficulty of hiring data scientists who are able to produce business value and work well with colleagues.

---

[2] I use the term *data scientist* as an umbrella term for statistics, operations research, business analytics, machine learning, etc. The term *scientist* nicely communicates the degree of innovation and uncertainty surrounding our work.

Business skills can make or break your career. In this book, I cover the skills that I've seen to be most critical for data scientists. This book is based partly on my own experience as a data scientist working across a wide range of countries and industries, partly on principles from professional publications, and partly on input I've received from the hundreds of data scientists who have attended my business skills training sessions in recent years.

The topics I cover in the following twelve chapters will have a significant impact on the success of your career. Each chapter introduces basic principles in one key area, but be aware that entire books have been written on the subjects of some of these individual chapters. I'll generally recommend additional references as appropriate. Think of this book, then, as an overview of key topics and a point of entry for further learning where you need it.

The six topics in this book—company, colleagues, communication (storytelling), expectations, results and careers—are covered in six sections. They are presented roughly in the order you'll need them during your career, with project leadership skills coming later in the book. The final topic, careers, will be relevant at all stages of your career, even as you pursue your first job as a data scientist.

The first section, **company**, will help you work within a larger organization. Chapter one discusses how organizations generally view data scientists, why they hire them and what they expect from them. This chapter begins to explain how you as a data scientist can produce business value, a topic that is more fully developed in chapter ten. Chapter two describes how to relate to the diverse range of individuals within your organization. By understanding how various roles and functions differ in their goals and values, you will be better prepared to deliver valuable results to them.

The second section, **colleagues**, covers interpersonal challenges that often present significant problems for data scientists—cultural

misunderstandings, negotiating through disagreements, and office politics. Younger data scientists generally perceive these topics as less important, until the day when they suddenly find them extremely important.

The third section, **storytelling**, covers principles and methods key to oral and visual presentations. This section describes how to clarify your message, which is easy in principle but difficult in practice, and how to produce clear slides, graphs and tables. It goes into detail on techniques for decluttering and drawing focus, leveraging principles such as Gestalt and preattentive processing. The two chapters in this section are perhaps the most immediately relevant for many data scientists and will provide you with some very quick wins, as I've seen from the responses of my training participants.

The fourth section, **expectations**, covers techniques for setting expectations, kicking off a project, and maintaining the trust of your stakeholders. In particular, I cover ways to build consensus at the start of a project and ways to avoid the common mistake of working hard to build a solution that no one wants to use.

The fifth section, **results**, presents principles, techniques, and case studies related to selecting the most beneficial data science projects. It then describes concepts and frameworks that will help you run those projects to successful completion, despite the shifting expectations and lack of clarity so prevalent in data science projects.

The sixth and final section, **careers**, addresses questions I'm often asked by data scientists, such as how to choose between job opportunities, how to strengthen your CV, and even whether to start working independently as a freelance data scientist.

I've resisted the urge to include definitions for terms such as *MVP, KPI, stakeholder,* etc., within the text, but such terms (indicated in bold at their first occurrence) are defined in a glossary at the end.

Despite the wide range of topics covered in this book, there is one central theme. The foundation for all skills I will discuss is **the skill of understanding the perspectives of people around you**, especially the people with whom you, as a data scientist, have the least in common. This foundational skill is often called *empathy*. Understanding the perspectives of the diverse range of colleagues you work with will help you relate more effectively in a business setting, ultimately enabling you to deliver successful data science projects with real business impact.