

1

Introduction

1.1 MODELS AND MODELING

Modeling is the process of creating a simplified representation of reality and working with this representation in order to understand or control some aspect of the world. While this book is devoted to *mathematical* models, modeling itself is a ubiquitous human activity. In fact, it seems to be one of just a few fundamental ways in which we humans understand our environment.

As an example, a map is one of the most common models we encounter. Maps are models because they simplify reality by leaving out most geographic details in order to highlight the important features we need. A state road map, for example, shows major roads but not minor ones, gives rough locations of cities but not individual addresses, and so on. The map we choose must be appropriate for the need we have: a long trip across several states requires a regional map, while a trip across town requires a detailed street map. In the same way, a good model must be appropriate for the specific uses intended for it. A complex model of the economy is probably not appropriate for pricing an individual product. Similarly, a back-of-the-envelope calculation is likely to be inappropriate for acquiring a multibillion-dollar company.

Models take many different forms: mental, visual, physical, mathematical, and spreadsheet, to name a few. We use mental models constantly to understand the world and to predict the outcomes of our actions. Mental models are informal, but they do allow us to make a quick judgment about the desirability of a particular proposal. For example, mental models come into play in a hiring decision. One manager has a mental model that suggests that hiring older workers is not a good idea because they are slow to adopt new ways; another manager has a mental model that suggests hiring older workers is a good idea because they bring valuable experience to the job. We are often unaware of our own mental models, yet they can have a strong influence on the actions we take, especially when they are the primary basis for decision making.

While everyone uses mental models, some people routinely use other kinds of models in their professional lives. Visual models include maps, as we mentioned earlier. Organization charts are also visual models. They may represent reporting relationships, reveal the locus of authority, suggest major channels of communication, and identify responsibility for personnel decisions. Visual models are used in various sports, for instance, as when a coach sketches the playing area and represents team members and opponents as *X*'s and *O*'s. Most players probably don't realize that they are using a model for the purposes of understanding and communication.

Physical models are used extensively in engineering to assist in the design of airplanes, ships, and buildings. They are also used in science, as, for example, in depicting the spatial arrangement of amino acids in the DNA helix or the makeup of a chemical compound. Architects use physical models to show how a proposed building fits within its surroundings.

Mathematical models take many forms and are used throughout science, engineering, and public policy. For instance, a groundwater model helps determine where flooding is most likely to occur, population models predict the spread of infectious disease, and exposure-assessment models forecast the impact of toxic spills. In other settings, traffic-flow models predict the buildup of highway congestion, fault-tree models help reveal the causes of an accident, and reliability models suggest when equipment may need

2 CHAPTER 1 INTRODUCTION

replacement. Mathematical models can be extremely powerful, especially when they give clear insights into the forces driving a particular outcome.

1.1.1 Why Study Modeling?

What are the benefits of building and using formal models, as opposed to relying on mental models or just “gut feel?” The primary purpose of modeling is to generate *insight*, by which we mean an improved understanding of the situation or problem at hand. While mathematical models consist of numbers and symbols, the real benefit of using them is to make better *decisions*. Better decisions are most often the result of improved understanding, not just the numbers themselves.

Thus, we study modeling primarily because it improves our thinking skills. Modeling is a discipline that provides a structure for problem solving. The fundamental elements of a model—such as parameters, decisions, and outcomes—are useful concepts in all problem solving. Modeling provides examples of clear and logical analysis and helps raise the level of our thinking.

Modeling also helps improve our quantitative reasoning skills. Building a model demands care with units and with orders of magnitude, and it teaches the importance of numeracy. Many people are cautious about quantitative analysis because they do not trust their own quantitative skills. In the best cases, a well-structured modeling experience can help such people overcome their fears, build solid quantitative skills, and improve their performance in a business world that demands (and rewards) these skills.

Any model is a laboratory in which we can experiment and learn. An effective modeler needs to develop an open, inquiring frame of mind to go along with the necessary technical skills. Just as a scientist uses the laboratory to test ideas, hypotheses, and theories, a business analyst can use a model to test the implications of alternative courses of action and develop not only a recommended decision but, equally important, the rationale for why that decision is preferred. The easy-to-understand rationale behind the recommendation often comes from insights the analyst has discovered while testing a model.

1.1.2 Models in Business

Given the widespread use of mathematical models in science and engineering, it is not surprising to find that they are also widely used in the business world. We refer to people who routinely build and analyze formal models in their professional lives as **business analysts**. In our years of training managers and management students, we have found that strong modeling skills are particularly important for consultants, as well as for financial analysts, marketing researchers, entrepreneurs, and others who face challenging business decisions of real economic consequence. Practicing business analysts and students intending to become business analysts are the intended audience for this book.

Just as there are many types of models in science, engineering, public policy, and other domains outside of business, many different types of models are used in business. We distinguish here four model types that exemplify different levels of interaction with, and participation by, the people who use the models:

- One-time decision models
- Decision-support models
- Models embedded in computer systems
- Models used in business education

Many of the models business analysts create are used in **one-time decision problems**. A corporate valuation model, for example, might be used intensively during merger negotiations but never thereafter. In other situations, a one-time model might be created to evaluate the profit impact of a promotion campaign, or to help select a health insurance provider, or to structure the terms of a supply contract. One-time models are usually built by decision makers themselves, frequently under time pressure. Managerial judgment is often used as a substitute for empirical data in such models, owing to time constraints and data limitations. Most importantly, this type of model involves the user intensively because the model is usually tailored to a particular decision-making need. One major benefit of studying modeling is to gain skills in building and using one-time models effectively.

Decision-support systems are computer systems that tie together models, data, analysis tools, and presentation tools into a single integrated package. These systems are intended for repeated use, either by executives themselves or by their analytic staff. Decision-support systems are used in research and development planning at pharmaceutical firms, pricing decisions at oil companies, and product-line profitability analysis at manufacturing firms, to cite just a few examples. Decision-support systems are usually built and maintained by information systems personnel, but they represent the routine use of what were once one-time decision models. After a one-time model becomes established, it can be adapted for broader and more frequent use in the organization. Thus, the models within decision-support systems may initially be developed by managers and business analysts, but later streamlined by information systems staff for a less intensive level of human interaction. An additional benefit of studying modeling is to recognize possible improvements in the design and operation of decision-support systems.

Embedded models are those contained within computer systems that perform routine, repeated tasks with little or no human involvement. Many inventory replenishment decisions are made by automated computer systems. Loan payments on auto leases or prices for stock options are also determined by automated systems. Routine real estate appraisals may also be largely automated. In these cases, the models themselves are somewhat hidden in the software. Many users of embedded models are not aware of the underlying models; they simply assume that the “system” knows how to make the right calculations. An ancillary benefit of studying modeling is to become more aware, and perhaps more questioning, of these embedded models.

1.1.3 Models in Business Education

Models are useful not only in the business world, but also in the academic world where business analysts are educated. The modern business curriculum is heavily dependent on models for delivering basic concepts as well as for providing numerical results. An introductory course in Finance might include an option-pricing model, a cash-management model, and the classic portfolio model. A basic Marketing course might include demand curves for pricing analysis, a diffusion model for new-product penetration, and clustering models for market segmentation. In Operations Management, we might encounter inventory models for stock control, allocation models for scheduling production, and newsvendor models for trading off shortage and surplus outcomes. Both micro- and macroeconomics are taught almost exclusively through models. Aggregate supply-and-demand curves are models, as are production functions.

Most of the models used in education are highly simplified, or *stylized*, in order to preserve clarity. Stylized models are frequently used to provide insight into qualitative phenomena, not necessarily to calculate precise numerical results. In this book, we frequently use models from business education as examples, so that we can combine learning about business with learning about models. In fact, the tools presented in this book can be used throughout the curriculum to better understand the various functional areas of business.

1.1.4 Benefits of Business Models

Modeling can benefit business decision making in a variety of ways.

- Modeling allows us to make inexpensive errors. Wind-tunnel tests are used in airplane design partly because if every potential wing design had to be built into a full-scale aircraft and flown by a pilot, we would lose far too many pilots. In a similar way, we can propose ideas and test them in a model, without having to suffer the consequences of bad ideas in the real world.
- Modeling allows us to explore the impossible. Many companies have policies, procedures, or habits that prevent them from making certain choices. Sometimes these habits prevent them from discovering better ways of doing business. Modeling can be used to explore these “impossible” alternatives and to help convince the skeptics to try a different approach.
- Modeling can improve business intuition. As we have said, a model is a laboratory in which we perform experiments. We can usually learn faster from laboratory experiments than from experience in the real world. With a model, we can try thousands of

4 CHAPTER 1 INTRODUCTION

combinations that would take many years to test in the real world. We can also try extreme ideas that would be too risky to test in the real world. And we can learn about how the world works by simulating a hundred years of experience in just a few seconds.

- Modeling provides information in a timely manner. For example, while a survey could be used to determine the potential demand for a product, effective modeling can often give useful bounds on the likely range of demand in far less time.
- Finally, modeling can reduce costs. Data collection is often expensive and time-consuming. An effective modeler may be able to provide the same level of information at a much lower cost.

Even among those who do not build models, skill in working with models is very important. Most business students eventually find themselves on a team charged with recommending a course of action. If these teams do not build models themselves, they often work with internal or external consultants who do. Experience in building and analyzing models is, in our minds, the best training for working effectively on problem-solving teams. People who have not actually built a few models themselves often accept model results blindly or become intimidated by the modeling process. A well-trained analyst not only appreciates the power of modeling but also remains skeptical of models as panaceas.

We believe that modeling skills are useful to a very broad range of businesspeople, from junior analysts without a business degree to senior vice presidents who do their own analysis. Many recent graduates have only a superficial knowledge of these tools because their education emphasized passive consumption of other people's models rather than active model building. Thus, there is considerable potential even among master's-level graduates to improve their modeling skills so that they can become more capable of carrying out independent analyses of important decisions. The only absolute prerequisite for using this book and enhancing that skill is a desire to use logical, analytic methods to reach a higher level of understanding in the decision-making world.

1.2 THE ROLE OF SPREADSHEETS

Because spreadsheets are the principal vehicle for modeling in business, spreadsheet models are the major type we deal with in this book. Spreadsheet models are also mathematical models, but, for many people, spreadsheet mathematics is more accessible than algebra or calculus. Spreadsheet models do have limitations, of course, but they allow us to build more detailed and more complex models than traditional mathematics allows. They also have the advantage of being pervasive in business analysis. Finally, the spreadsheet format corresponds nicely to the form of accounting statements that are used for business communication; in fact, the word "spreadsheet" originates in accounting and only recently has come to mean the electronic spreadsheet.

It has been said that the spreadsheet is the *second best* way to do many kinds of analysis and is therefore the *best* way to do most modeling. In other words, for any one modeling task, a more powerful, flexible, and sophisticated software tool is almost certainly available. In this sense, the spreadsheet is the Swiss Army knife of business analysis. Most business analysts lack the time, money, and knowledge to learn and use a different software tool for each problem that arises, just as most of us cannot afford to carry around a complete toolbox to handle the occasional screw we need to tighten. The practical alternative is to use the spreadsheet (and occasionally one of its sophisticated add-ins) to perform most modeling tasks. An effective modeler will, of course, have a sense for the limitations of a spreadsheet and will know when to use a more powerful tool.

Despite its limitations, the electronic spreadsheet represents a breakthrough technology for practical modeling. Prior to the 1980s, modeling was performed only by specialists using demanding software on expensive hardware. This meant that only the most critical business problems could be analyzed using models because only these problems justified the large budgets and long time commitments required to build, debug, and apply the models of the day. This situation has changed dramatically in the past 30 years or so. First the personal computer, then the spreadsheet, and recently the arrival of add-ins for specialized analyses have put tremendous analytical power at the hands of

anyone who can afford a laptop and some training. In fact, we believe the 1990s will come to be seen as the dawn of the “end-user modeling” era. End-user modelers are analysts who are not specialists in modeling, but who can create an effective spreadsheet and manipulate it for insight. The problems that end-user modelers can solve are typically not the multibillion-dollar, multiyear variety; those are still the preserve of functional-area specialists and sophisticated computer scientists. Rather, the end user can apply modeling effectively to hundreds of important but smaller-scale situations that in the past would not have benefited from this approach. We provide many illustrations throughout this book.

Spreadsheet skills themselves are now in high demand in many jobs, although experts in Excel may not be skilled modelers. In our recent survey of MBAs from the Tuck School of Business (available at <http://mba.tuck.dartmouth.edu/spreadsheet/>), we found that 77 percent said that spreadsheets were either “very important” or “critical” in their work. Good training in spreadsheet modeling, in what we call **spreadsheet engineering**, is valuable because it can dramatically improve both the efficiency and effectiveness with which the analyst uses spreadsheets.

1.2.1 Risks of Spreadsheet Use

Countless companies and individuals rely on spreadsheets every day. Most users assume their spreadsheet models are error free. However, the available evidence suggests just the opposite: many, perhaps most, spreadsheets contain internal errors, and more errors are introduced as these spreadsheets are used and modified. Given this evidence, and the tremendous risks of relying on flawed spreadsheet models, it is critically important to learn how to create spreadsheets that are as close to error free as possible and to use spreadsheets in a disciplined way to avoid mistakes.

It is rare to read press reports on problems arising from erroneous spreadsheets. Most companies do not readily admit to these kinds of mistakes. However, the few reports that have surfaced are instructive. For many years the European Spreadsheet Risks Interest Group (EUSPRIG) has maintained a website (<http://www.eusprig.org/horror-stories.htm>) that documents dozens of verified stories about spreadsheet errors that have had a quantifiable impact on the organization. Here is just a small sample:

- Some candidates for police officer jobs are told that they have passed the test when in fact they have failed. Reason: improper sorting of the spreadsheet.
- An energy company overcharges consumers between \$200 million and \$1 billion. Reason: careless naming of spreadsheet files.
- A think-tank reports that only 11 percent of a local population has at least a bachelor’s degree when in fact the figure is 20 percent. Reason: a copy-and-paste error in a spreadsheet.
- Misstated earnings lead the stock price of an online retailer to fall 25 percent in a day and the CEO to resign. Reason: a single erroneous numerical input in a spreadsheet.
- A school loses £30,000 because its budget is underestimated. Reason: numbers entered as text in a spreadsheet.
- The Business Council reports that its members forecast slow growth for the coming year when their outlook is actually quite optimistic. Reason: the spreadsheet shifted, so the wrong numbers appeared in the wrong columns.
- Benefits of unbundling telecommunication services are understated by \$50 million. Reason: incorrect references in a spreadsheet formula.

These cases suggest that spreadsheets can lead to costly errors in a variety of ways. But are spreadsheets themselves properly *built* in the first place? Apparently they are not, at least according to several research studies. In our own investigation of 50 spreadsheets that were being used by various organizations, fewer than 10 percent were free of errors.¹ This evidence serves notice that errors in spreadsheets may be rampant and insidious.

Despite the research evidence, very few corporations employ even the most basic design methodologies and error-detection procedures. These procedures take time and effort, whereas one of the great appeals of spreadsheet modeling is that it can be done

¹ S. Powell, K. Baker and B. Lawson, “Errors in Operational Spreadsheets,” *Journal of End User Computing* 21, (July–September, 2009): 24–36.

6 CHAPTER 1 INTRODUCTION

quickly and easily, even by business analysts who are not professional programmers. But ease of use is a delusion if the results contain significant errors.

Briefly stated, the business world is still at an early stage of understanding how to develop error-free spreadsheets. Organizations are in need of better methods for detecting errors and more reliable procedures for preventing errors in the first place. However, the research literature on these topics has not advanced very far, and the state of the art remains somewhat primitive.

1.2.2 Challenges for Spreadsheet Users

Spreadsheets represent the ubiquitous software platform of business. Millions of spreadsheet models are used each day to make decisions involving billions of dollars, and thousands of new spreadsheets come into being each day. Given this usage pattern, we might think that spreadsheet engineering is a well-developed discipline and that expertise in spreadsheet modeling can be found in just about any company. Amazingly, the opposite is true.

What is the current state of spreadsheet use by end-user modelers? The evidence available from audits of existing spreadsheets, laboratory experiments, surveys of end users, and field visits suggests that, despite widespread use, the quality with which spreadsheets are engineered generally remains poor. There are four major problem areas:

- End-user spreadsheets frequently have bugs.
- End users are overconfident about the quality of their own spreadsheets.
- The process that end users employ to create their spreadsheets is inefficient at best and chaotic at worst.
- End users fail to employ the most productive methods for generating insights from their spreadsheets.

Our own research, conducted as part of the Spreadsheet Engineering Research Project (<http://mba.tuck.dartmouth.edu/spreadsheet/>), found that a substantial majority of spreadsheets in use contain at least one error. A follow-up study found that most of these errors had a substantial impact on the quantitative analysis in the spreadsheets. However, our investigation also suggested that errors in individual cells may be only a symptom. The underlying cause often seems to be a high degree of complexity in the model, even when the corresponding problem is relatively simple. Complexity arises in many ways:

- Individual cell formulas that are excessively long and involved
- Poorly designed worksheets that are difficult to navigate and understand
- Poorly organized workbooks whose underlying structure is concealed

Spreadsheets that are overly complex and difficult for anyone other than the designer to use, even if they are technically correct, may be the cause of some of the costly mistakes attributed to spreadsheets.

Laboratory experiments have uncovered another disturbing fact about spreadsheet modeling: end users appear to be overconfident about the likelihood of errors in their own spreadsheets. In these experiments, undergraduate volunteers were asked to build a spreadsheet for a well-defined problem. After they were done, the volunteers were given time to review and audit their models. Finally, they were asked to evaluate the likelihood that their model contained one or more bugs. While 18 percent of the subjects thought their models had one or more bugs, the actual proportion proved to be 80 percent. That is, 80 percent of these spreadsheets actually had bugs, but only about 18 percent of those who built them suspected they had bugs. This finding of overconfidence is consistent with the findings of other studies: people tend to underestimate the possibility that they might make mistakes. Unfortunately, this overconfidence translates directly into a casual attitude toward spreadsheet design and ultimately into a disturbingly high error rate among spreadsheets in actual use.

Our observations and research into how end users actually construct spreadsheets suggest that the process is often inefficient:

- End users typically do not plan their spreadsheets. Instead, they build them live at the keyboard. The result in many cases is extensive rework. (In our survey of MBA graduates, we found that about 20 percent sketched a spreadsheet on paper first,

whereas about 50 percent started by entering data and formulas directly into the computer.)

- End users do not use a conscious prototyping approach, which involves building a series of models starting with the simplest and gradually adding complexity.
- End users rarely spend time debugging their models, unless the model performs in such a counterintuitive manner that it demands intervention.
- End users almost never subject their spreadsheets to review by another person. In general, end users appear to trust that the model they *thought* they had built is actually the model they *see* on their screens, despite the fact that spreadsheets show only numbers, not the relationships behind the numbers.
- Finally, many end users, even some who are experts in Excel, do not consistently use tools that can help generate the insights that make modeling worthwhile. Excel's Data Table and Goal Seek tools, to cite just two examples, are overlooked by the majority of end users. Without these tools, the end user either fails to ask questions that can provide telling insights, or else wastes time generating results that could be found more easily.

The evidence is strong that the existing state of spreadsheet design and use is generally inadequate. This is one reason we devote a significant portion of this book to spreadsheet engineering. Only with a solid foundation in spreadsheet engineering can the business analyst effectively generate real insights from spreadsheet models.

1.2.3 Background Knowledge for Spreadsheet Modeling

Many people new to modeling fear it because modeling reminds them of painful experiences with mathematics. We do not wish to downplay the essentially mathematical nature of modeling, even modeling using spreadsheets. However, an effective modeler does not need to know any really advanced math. Knowledge of basic algebra (including functions such as the quadratic, exponential, and logarithmic), simple logic (as expressed in an IF statement or the MAX function), and basic probability (distributions and sampling, for example) will usually suffice. When we find it necessary to use any higher math in this book, we provide explanations. But our focus here is less on the mathematical details of models than on the creative process of constructing and using models.

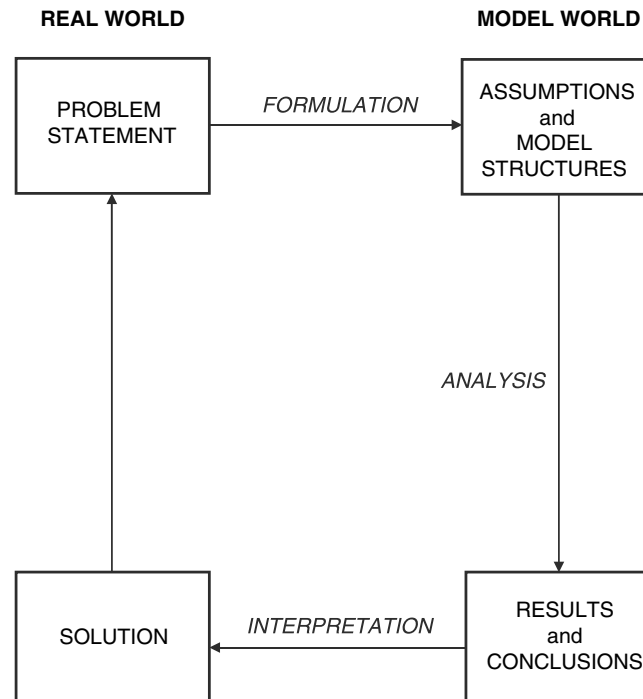
We assume throughout this book that the reader has a basic familiarity with Excel. This includes the ability to build a simple spreadsheet, enter and format text and data, use formulas and simple functions such as SUM, construct graphs, and so on. We do not assume the reader is an expert in Excel, nor do we assume knowledge of the advanced tools we cover, such as optimization and simulation. We have found that, in many situations, advanced Excel skills are not required for building effective models. And we believe that the main purpose of modeling is to improve the insight of the modeler. Thus, it is appropriate for a modeler with only basic Excel skills to build a model using only basic tools, and it is appropriate for a modeler with advanced skills to draw on advanced tools when needed. We have also found that too much skill in Excel can sometimes distract from the essential modeling tasks, which are almost always more about finding a simple and effective representation of the problem at hand than about finding some Excel trick.

For easy reference, we have included Appendix 1 to give an overview of Excel, from the basics of entering text and data to advanced formulas and functions. In addition, Appendix 2 covers the use of macros and an introduction to Visual Basic for Applications (VBA). We expect most readers to already know Excel to some degree, and to use these appendices as needed to hone specific skills. We believe that, by working through the examples in the book, the reader's Excel skills will improve naturally and painlessly, just as ours have improved over years of building models and teaching modeling to students whose Excel skills often exceeded ours.

1.3 THE REAL WORLD AND THE MODEL WORLD

We stated at the outset that modeling provides a structure for problem solving. It does this through a process of abstraction, in which the essence of the problem is captured in a simplified form. Because of this abstraction process, modeling does not come naturally to

8 CHAPTER 1 INTRODUCTION

FIGURE 1.1 The Real-World and the Model World

most people but must be learned. Because it does not come naturally, it can appear to be artificial and counterintuitive, causing many students of modeling to become uncomfortable with the process. This section attempts to reduce that discomfort by placing modeling in the context of problem solving in the real world.

A model is an abstraction, or simplification, of the real world. It is a laboratory—an artificial environment—in which we can experiment and test ideas without the costs and risks of experimenting with real systems and organizations. Figure 1.1 is a schematic showing how modeling creates an artificial world. We begin in the real world, usually with a messy problem to solve. If we determine that modeling is an appropriate tool, we then move across an invisible boundary into the model world.

In order to move into the model world, we abstract the essential features of the real world, leaving behind all the nonessential detail and complexity. We then construct our laboratory by combining our abstractions with specific assumptions and building a model of the essential aspects of the real world. This is the process of **model formulation**. It is an exercise in simplifying the actual situation and capturing its essence, with a specific purpose in mind. The model formulation process typically forces us to confront four features of a model:

- Decisions
- Outcomes
- Structure
- Data

Decisions refers to possible choices, or courses of action, that we might take. These would be controllable variables, such as quantities to buy, manufacture, spend, or sell. (By contrast, uncontrollable variables such as tax rates or the cost of materials are not decision variables.) **Outcomes** refers to the consequences of the decisions—the performance measures we use to evaluate the results of taking action. Examples might include profit, cost, or efficiency. **Structure** refers to the logic and the mathematics that link the elements of our model together. A simple example might be the equation $P = R - C$, in which profit is calculated as the difference between revenue and cost. Another example might be the relationship $F = I + P - S$, in which final inventory is calculated from initial inventory, production, and shipments. Finally, **data** refers to specific numerical assumptions. That may mean actual observations of the real world (often called “raw” or “empirical” data), or it may mean estimates of uncontrollable variables in the problem’s environment. Examples might include the interest rate on borrowed funds, the production capacity of a manufacturing facility, or the first-quarter sales for a new product.

Once it is built, we can use the model to test ideas and evaluate solutions. This is a process of **analysis**, in which we apply logic, often with the support of software, to take us from our assumptions and abstractions to a set of derived conclusions. Unlike model formulation, which tends to be mostly an art, analysis is much more of a science. It relies on mathematics and reason in order to explore the implications of our assumptions. This exploration process leads, hopefully, to insights about the problem confronting us. Sometimes, these insights involve an understanding of why one solution is beneficial and another is not; at other times, the insights involve understanding the sources of risk in a particular solution. In another situation, the insights involve identifying the decisions that are most critical to a good result, or identifying the inputs that have the strongest influence on a particular outcome. In each instance, it is crucial to understand that these insights are derived from the *model* world and not from the *real* world. Whether they apply to the real world is another matter entirely and requires managerial judgment.

To make the model insights useful, we must first translate them into the terms of the real world and then communicate them to the actual decision makers involved. Only then do *model* insights turn into useful *managerial* insights. And only then can we begin the process of evaluating solutions in terms of their impact on the real world. This is a process of **interpretation**, and here again, the process is an art. Good modelers can move smoothly back and forth between the model world and the real world, deriving crisp insights from the model, and translating the insights, modifying them as needed, to account for real-world complexities not captured in the model world.

This schematic description of the modeling process highlights some of the reasons it can be a challenge to incorporate modeling into problem solving. Powerful in competent hands, modeling is also somewhat esoteric. It involves deliberate abstraction and simplification of a situation, which appears to many people as a counterproductive exercise. Modeling requires a willingness to temporarily set aside much of the richness of the real world and to operate in the refined and artificial world of models and model insights. It also requires confidence that whatever insights arise in the model world can be translated into useful ideas in the real world. In addition, it requires an ability to mix art with science in order to exploit the modeling process to its full potential. Until we have some experience with this process, we may be resistant and skeptical. And it is always easy to criticize a model as being too simple. Good models are as simple as they can possibly be. But this very simplicity can appear to be a fatal flaw to skeptics. Nevertheless, modeling is one of the most powerful tools in the problem solver's tool kit, simply because there is no more practical way to arrive at the insights modeling can provide.

1.4 LESSONS FROM EXPERT AND NOVICE MODELERS

Perhaps the best way to become a good modeler is to serve an apprenticeship under an expert. Unfortunately, such opportunities are rare. Moreover, experts in all fields find it difficult to express their expertise or to teach it. While narrow, **technical skills** are relatively easy to teach (e.g., how to use the NPV function in Excel), expertise consists largely of **craft skills** that are more difficult to teach (e.g., what to include and exclude from the model). In the arts, there is a tradition of studio training, where a teacher poses artistic challenges to students and then coaches them as they work through the problems on their own. This is one way for students to acquire some of the difficult-to-articulate craft skills of the master. There is no comparable tradition in the mathematical fields; in fact, there is a long-standing belief that modeling cannot be taught but must simply be acquired by experience.

One way to improve modeling skills is to understand what expert and novice modelers actually do when they build and use models. From closely observing experts, we can attempt to articulate a set of modeling best practices. From observing novices, we can understand the reasons for their relatively lower level of modeling accomplishment: the blind alleys, counterproductive behaviors, misperceptions, and cognitive limitations that keep them from attaining expert performance. In this section, we summarize research studies on both expert and novice modelers.

1.4.1 Expert Modelers

An alternative to an apprenticeship under an expert is to study experts in a laboratory setting. Tom Willemain did this in a series of experiments with 12 expert modelers.

10 CHAPTER 1 INTRODUCTION

He gave each expert a short problem description as it would come from a client and observed the subject working for one hour on the problem. The subjects were asked to think out loud so that their thought processes could be recorded. Willemain's results concerning the "first hour in the life of a model" are highly suggestive of some of the ingredients of good modeling practice.²

Willemain was interested in determining the issues to which expert modelers devote attention as they formulate their models. He identified five topics important to modelers:

- Problem context
- Model structure
- Model realization
- Model assessment
- Model implementation

Problem context refers to the situation from which the modeler's problem arises, including the client, the client's view of the problem, and any available facts about the problem. In this activity, the modeler tries to understand the problem statement as provided by the client and to understand the messy situation out of which the problem arises.

Model structure refers to actually building the model itself, including issues such as what type of model to use, where to break the problem into subproblems, and how to choose parameters and relationships. In Figure 1.1, this would be the process of moving into the model world, making abstractions and assumptions, and creating an actual model.

Model realization refers to the more detailed activities of fitting the model to available data and calculating results. Here, the focus is on whether the general model structure can actually be implemented with the available data and whether the type of model under development will generate the hoped-for kinds of results. This topic corresponds to the analysis process in Figure 1.1.

Model assessment includes evaluating the model's correctness, feasibility, and acceptability to the client. Determining the correctness of a model involves finding whether the model assumptions correspond well enough to reality. Feasibility refers to whether the client has the resources to implement the developed model, whether sufficient data are available, and whether the model itself will perform as desired. Client acceptability refers to whether the client will understand the model and its results and whether the results will be useful to the client. In this phase, we can imagine the modeler looking from the model world back into the real world and trying to anticipate whether the model under construction will meet the needs of the client.

Finally, **model implementation** refers to working with the client to derive value from the model. This corresponds to the interpretation activity in Figure 1.1.

One of Willemain's interesting observations about his experts was that they frequently switched their attention among these five topics. That is, they did not follow a sequential problem-solving process, but rather moved quickly among the various phases—at one moment considering the problem statement, at another considering whether the necessary data would be available, and at yet another thinking through whether the client could understand and use the model. A second significant finding was that model structure, presumably the heart of a modeler's work, received a relatively small amount of attention (about 60 percent of the effort) when compared to the other four topics. Finally, it turned out that experts often alternated their attention between model structure and model assessment. That is, they would propose some element of model structure and quickly turn to evaluating its impact on model correctness, feasibility, and acceptability. Willemain suggests that the experts treat model structuring as the central task, or backbone, of their work, but they often branch off to examine related issues (data availability, client acceptance, and so on), eventually returning to the central task. In effect, model structuring becomes an organizing principle, or mental focus, around which the related activities can be arrayed.

The overall picture that emerges from this research is one in which craft skills are as essential to the effective modeler as technical skills. An effective modeler must understand the problem context, including the client, or modeling will fail. Similarly, a model that is

² T.R. Willemain, "Insights on Modeling from a Dozen Experts," *Operations Research* 42, No. 2 (1994): 213–222; "Model Formulation: What Experts Think About and When," *Operations Research* 43, No. 6 (1995): 916–932.

technically correct but does not provide information the client can use, or does not gain the trust of the client, represents only wasted effort. Experts approach modeling with a general process in mind, but they move fairly quickly among the different activities, creating, testing, and revising constantly as they go. The experts appear to be comfortable with a high degree of ambiguity as they approach the task of structuring a model. They do not rush to a solution, but patiently build tentative models and test them, always being ready to revise and improve.

1.4.2 Novice Modelers

Novices have been studied in many domains, from solving physics problems to playing golf. In general, novice problem solvers can be expected to show certain kinds of counterproductive behaviors. One is that they focus on just one approach to a problem and devote all their time to it, while experts are likely to try many different approaches. Novices also do not evaluate their performance as frequently or as critically as expert problem solvers do. Finally, novices tend to attempt to solve a problem using only the information given in that problem, while experts are more likely to draw on experience with other problems for useful analogies or tools.

In an attempt to better understand how our own students model problems, we conducted an experiment similar in most respects to Willemain's experiment with experts.³ We audiotaped 28 MBA students while they worked through four ill-structured modeling problems. Thus, this experiment did not focus on building a spreadsheet model for a well-defined problem, as might be assigned in a course for homework, but rather on formulating an approach to an ill-structured problem of the kind that consultants typically encounter. (Some of these problems will be presented in Chapter 2.) The students were given 30 minutes to work on each problem. The task was to begin developing a model that could ultimately be used for forecasting or for analysis of a decision.

We observed five behaviors in our subjects that are not typical of experts and that limit their modeling effectiveness:

- Overreliance on given numerical data
- Use of shortcuts to an answer
- Insufficient use of abstract variables and relationships
- Ineffective self-regulation
- Overuse of brainstorming relative to structured problem solving

In the study, some of the problems included extensive tables of numerical data. In these problems, many subjects devoted their time to examining the data rather than building a general model structure. Having data at hand seemed to block these students from the abstraction process required for effective modeling. In other problems, very little data was provided, and in these cases, some students attempted to "solve" the problem by performing calculations on the given numbers. Again, the data seemed to block the abstraction process. Many subjects complained about the lack of data in problems in which little was given, seeming to believe that data alone could lead to a solution. In general, then, our subjects appear to rely more on data than do experts, who build general model structures and only tangentially ask whether data exist or could be acquired to refine or operationalize their model structures.

Another problematic behavior we observed in our subjects was taking a shortcut to an answer. Where experts would consider various aspects of a problem and try out several different approaches, some students rushed to a conclusion. Some would simply rely on intuition to decide that the proposal they were to evaluate was a good or bad idea. Others would use back-of-the-envelope calculations to come to a conclusion. Still others would claim that the answer could be found by collecting data, or performing marketing research, or asking experts in the industry. (We call this behavior "invoking a magic wand.") All of these approaches seem to avoid the assigned task, which was to structure a *model* for analyzing the problem, not to come to a conclusion.

³ S.G. Powell and T.R. Willemain, "How Novices Formulate Models. Part I: Qualitative Insights and Implications for Teaching," *Journal of the Operational Research Society*, 58 (2007): 983–995; T.R. Willemain and S.G. Powell, "How Novices Formulate Models. Part II: A Quantitative Description of Behavior," *Journal of the Operational Research Society*, 58 (2007): 1271–1283.

12 CHAPTER 1 INTRODUCTION

Expert problem solvers generally use abstract variables and relationships in the early stages of modeling a problem. We saw very little of this in our subjects, who appeared to think predominantly in concrete terms, often using specific numbers. Expert modelers tend to be well trained in formal mathematics, and they naturally think in terms of variables and relationships. Our subjects were generally less well trained in mathematics but tended to have extensive experience with spreadsheets. Their approach to spreadsheet modeling involved minimal abstraction and maximal reliance on numbers. Our subjects did not often write down variables and functions, but they fairly often sketched or talked about a spreadsheet in terms of its row and column headings.

As we noted earlier, experts pause frequently during problem solving to evaluate the approach they are taking. They are also willing to try another approach if the current one seems unproductive. By contrast, many of our subjects did little self-evaluation during the experiment. Some focused more on the problem we had given them as a business problem than a modeling problem. So the special features that a model brings to analyzing a situation seemed lost on them. Without a clear goal, a typical subject would launch into a discussion of all the factors that might conceivably influence the problem. Only rarely did we observe a subject stopping and asking whether they were making progress toward a *model*.

Finally, the predominant problem-solving strategy we observed our subjects using could be described as unstructured problem exploration. For example, they would list issues in a rambling and unstructured manner, as if they were brainstorming, without attempting to organize their thoughts in a form that would support modeling. Structured problem solving, as used by experts, seeks to impose an organized plan on the modeling process.

In general our subjects failed to think in modeling terms—that is, by deciding what the outcome of the modeling process was to be and working backwards through variables and assumptions and relationships to the beginning. Instead, they explored a variety of (usually) unrelated aspects of the problem in a discursive manner.

What can a business analyst who wants to improve modeling skills learn from this research? First, expertise takes time and practice to acquire, and the novice should not expect to perform like an expert overnight. However, some expert behaviors are worth imitating from the start. Don't look for quick answers to the problem at hand, and don't expect the data to answer the problem for you. Rather, use what you know to build a logical structure of relationships. Use whatever language you are most comfortable with (algebra, a spreadsheet, and a sketch), but work to develop your ability to abstract the essential features of the situation from the details and the numbers. Keep an open mind, try different approaches, and evaluate your work often. Most important, look for opportunities to use modeling, and constantly upgrade both your technical and craft skills.

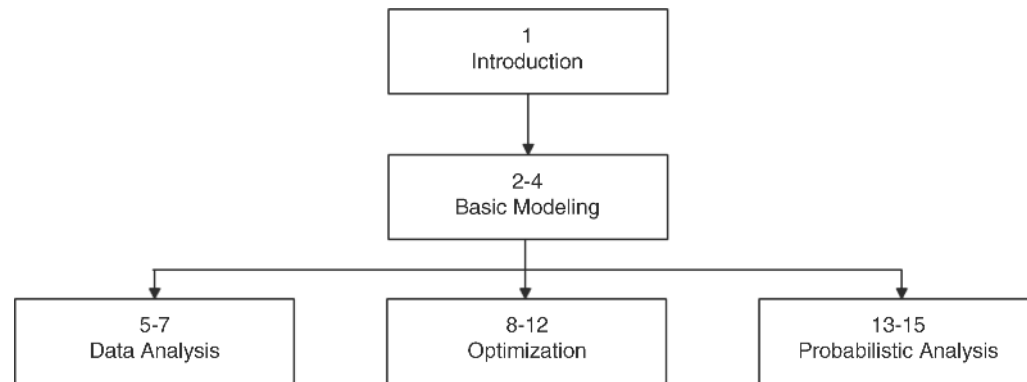
1.5 ORGANIZATION OF THE BOOK

This book is organized around the four sets of skills we believe business analysts most need in their modeling work:

- Spreadsheet engineering
- Modeling craft
- Data analysis
- Management science

Spreadsheet engineering deals with how to design, build, test, and perform analysis with a spreadsheet model. **Modeling craft** refers to the nontechnical but critical skills that an expert modeler employs, such as abstracting the essential features of a situation in a model, debugging a model effectively, and translating model results into managerial insights. **Data analysis** involves the exploration of datasets and the basic techniques used for classification and prediction. **Management science** covers optimization and simulation. A basic knowledge of these tools is important for the well-rounded analyst. Figure 1.2 provides an overview of the organization of the book.

The heart of this book is the material on building spreadsheet models and using them to analyze decisions. However, before the analyst can build spreadsheet models successfully, certain broader skills are needed. Therefore, we begin in Chapter 2 with a

FIGURE 1.2 Organization of the Book

discussion of the various contexts in which modeling is carried out and the role that modeling plays in a structured problem-solving process. We also introduce in this chapter the craft aspects of modeling—the tricks of the trade that experienced and successful modelers employ. These are not Excel tricks, but rather approaches to dealing with the ambiguities of analysis using models. Chapters 3 and 4 provide the essential tools of spreadsheet engineering. Along with the earlier material, these chapters should be studied by all readers. (Appendix 1 contains a brief overview of the Excel skills needed by effective modelers, and Appendix 2 provides a glimpse of the advanced capabilities available with Visual Basic for Applications.) Chapter 3 provides guidelines for designing effective spreadsheets and workbooks, while Chapter 4 provides an overview of various tools available for analyzing spreadsheet models. Chapters 5 through 15 cover the advanced tools of the management scientist and their spreadsheet implementations. Chapters 5 through 7 deal with data exploration, basic data mining, and forecasting. Chapters 8 through 12 explore optimization, and Chapters 13 through 15 cover simulation and probability-based models. (The necessary statistical background for our coverage appears in Appendix 3.) Numerous examples throughout the text illustrate good modeling techniques, and most chapters contain exercises for practice. Many of these exercises relate to a set of case problems, which are included at the end of the book. These problems provide an opportunity to gain experience with realistic modeling problems that build on concepts in different chapters.

1.6 SUMMARY

The following statements summarize the principles on which this book is based.

- *Modeling is a necessary skill for every business analyst.* Models are encountered frequently in business education and in the business world. Furthermore, analysts are capable of formulating their own models.
- *Spreadsheets are the modeling platform of choice.* The wide acceptance and flexibility of the spreadsheet make it the modeling platform of choice for most business situations. Since familiarity with spreadsheets is required for almost everyone in business, the basis for learning spreadsheet-based modeling is already in place.
- *Basic spreadsheet modeling skills are an essential foundation.* While basic knowledge about spreadsheets is usually assumed in business, spreadsheet skills and spreadsheet modeling skills are not the same. Effective education in business modeling begins with training in how to use a spreadsheet to build and analyze models.

- *End-user modeling is cost-effective.*

In an ever-growing range of situations, well-trained business analysts can build their own models without relying on consultants or experts.

- *Craft skills are essential to the effective modeler.*

Craft skills are the mark of an expert in any field. The craft skills of modeling must gradually be refined through experience, but the process can be expedited by identifying and discussing them and by providing opportunities to practice their use.

- *Analysts can learn the required modeling skills.*

Modeling skills do not involve complex mathematics or arcane concepts. Any motivated analyst can learn the basics of good modeling and apply this knowledge on the job.

- *Management science and data analysis are important advanced tools.*

Extensive knowledge of these tools is not required of most business analysts; however, solid knowledge of the fundamentals can turn an average modeler into a power modeler.

14 CHAPTER 1 INTRODUCTION

SUGGESTED READINGS

Many books are available on Excel, although most of them cover its vast array of features without isolating those of particular relevance for the business analyst. In the chapters on Excel, we provide several references to books and other materials for learning basic Excel skills. A working business analyst should probably own at least one Excel guide as a reference book. Two such references are:

Frye, C. 2015. *Microsoft Excel 2016 Step by Step*. Redmond, WA: Microsoft Press.
Walkenbach, J. 2015. *Excel 2016 Bible*, Indianapolis: Wiley Publishing.

Several textbooks present the tools of management science using spreadsheets. We recommend these for a more detailed treatment of management science than we provide here:

Albright, S. C. and W. Winston. 2015. *Business Analytics: Data Analysis and Decision Making*, 5th ed. Stamford, CT: Cengage Learning.
Ragsdale, C. 2012. *Spreadsheet Modeling and Decision Analysis*, 7th ed. Stamford, CT: Cengage Learning.

The standard reference on the mathematics of management science is:

Hillier, F., and G. Lieberman. 2009. *Introduction to Operations Research*, 9th ed. Oakland, CA: McGraw-Hill.

While this text does not rely on spreadsheets, it does provide in a relatively accessible form the methods behind much of the management science we present in this book. The following two references are more narrowly focused books that apply spreadsheet modeling to specific business disciplines:

Benninga, S. 2014. *Financial Modeling*, 4th ed. Cambridge, MA: MIT Press.
Lilien, G., and A. Rangaswamy. 2006. *Marketing Engineering*, 2nd ed. State College, PA: Decision Pro.

Finally, for stimulating books on modeling and problem solving, we recommend:

Casti, J. 1997. *Would-be Worlds: How Simulation Is Changing the Frontiers of Science*. New York: John Wiley & Sons.
Koomey, J. D. 2008. *Turning Numbers into Knowledge: Mastering the Art of Problem Solving*, 2nd ed. Oakland, CA: Analytics Press.
Starfield, A., K. Smith, and A. Bleloch. 1994. *How to Model It*. New York: McGraw-Hill.