Managerial Analytics

An Applied Guide to Principles, Methods, Tools, and Best Practices

Michael Watson | Derek Nelson

Praise for *Managerial Analytics*

"As the new era with abundant data sources is upon us, books on data analytics are gaining traction. Most of them either focus on the technical side of analytics or on the underlying business processes and the business of analytics. A manager who wants to learn the underlying techniques and their applicability is bound to tutorials from a data scientist with good communication skills. This book gives managers the opportunity to learn the concepts by themselves, and thus, it should be on a bookshelf of everyone who leads and manages analytics efforts. The book covers the most important methodologies and concepts in data analytics from a non-technical perspective. Each methodology is nicely wrapped with examples and use cases, and it does not require technical knowledge. If I were to venture in the field of analytics from the business perspective, this would be the first book to read in the morning."

—**Diego Klabjan,** Professor of Industrial Engineering and Management Sciences, Director of Master of Science in Analytics Program, Northwestern University

"An excellent introduction and overview of the field of analytics, *Managerial Analytics* is easily accessible for those new to the field and provides a useful framework for readers with a deeper background. Written from a practitioner's point of view, the book is well stocked with concise and relevant examples. The authors set out to define the boundaries of what is currently possible with analytics tools, and they guide the reader in asking good questions, avoiding common pitfalls, and identifying hidden assumptions when managing their own analytics projects."

—**Michael Freimer, Ph.D.,** Chief Scientist, DemandSignal

"The term 'analytics' means multiple things to different audiences. Watson and Nelson help to bridge the gap between the marketing hype and the technical details, making it easier to evaluate analytics-based solutions and better understand their potential."

—**Irv Lustig, Ph.D.,** Manager, Optimization and Mathematical Software, Business Analytics and Math Science, IBM Research at IBM

This page intentionally left blank

Managerial Analytics

This page intentionally left blank

Managerial Analytics

An Applied Guide to Principles, Methods, Tools, and Best Practices

Michael Watson
Derek Nelson

Vice President, Publisher: Tim Moore
Associate Publisher and Director of Marketing: Amy Neidlinger
Executive Editor: Jeanne Glasser Levine
Operations Manager: Jodi Kemper
Cover Designer: Alan Clements
Managing Editor: Kristy Hart
Project Editor: Elaine Wiley
Copy Editor: Kitty Wilson
Proofreader: Jess DeGabriele
Indexer: Lisa Stumpf
Senior Compositor: Gloria Schurick
Manufacturing Buyer: Dan Uhrig

Publishing as Pearson
Upper Saddle River, New Jersey 07458

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact international@pearsoned.com.

Printed in the United States of America

First Printing December 2013

ISBN-10: 0-13-340742-X
ISBN-13: 978-0-13-340742-6

Pearson Education LTD.
Pearson Education Australia PTY, Limited.
Pearson Education Singapore, Pte. Ltd.
Pearson Education Asia, Ltd.
Pearson Education Canada, Ltd.
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education—Japan
Pearson Education Malaysia, Pte. Ltd.

Library of Congress Control Number: 2013950966

To my wife, Kristen,
for all her support throughout this project
MSW

To my wife, Bridget,
who dreamed as a little girl of having a book about analytics dedicated to her ☺
DKN

This page intentionally left blank

Contents

Acknowledgments

This book, covering such a wide subject matter, would not have been possible without the many people we've learned so much from over the years.

While it is impossible to thank everyone, we apologize in advance if we've missed someone.

First, we would like to thank Pete Cacioppi for his help with early versions of the book and suggested text for certain optimization sections. Thanks also to Sara Lewis for reviewing and commenting on the pre-published book.

We'd like to thank all the experts who answered our questions and helped teach us about their specialty. Thank you Irv Lustig (for teaching us how to teach optimization and helping us understand the breadth of analytics), Richard Whisner and Patricio Cofre (for teaching us data modeling and BI systems), Dave Vander Veen (for teaching us about implementing full-scale analytics systems), David and Edith Simchi-Levi (for teaching us about rigorous analysis and how to commercialize this technology), Mark Daskin (for teaching us about optimization and how to present analytics results) and Diego Klabjan, the founding director of Northwestern University's Master of Science in Analytics, (for introducing us to many new ideas in the field of analytics programming).

We would like to thank the leading experts in the field who helped create the analytics movement. We've included an unofficial bibliography with some of the great books and articles we read. We also reference a lot of nice articles in the endnotes. And, we'll keep this online at the book's website (ManagerialAnalytics.com). Hopefully, this list will allow you to further explore topics of interest to you.

Finally, we couldn't have completed the book without the assistance and expert guidance of Jeanne Glasser Levine, our editor at Pearson.

About the Authors

Michael Watson is currently a partner at Opex Analytics and an Adjunct Professor at Northwestern University. At Opex Analytics he helps bring new analytics solutions to companies. Prior to Opex Analytics, he was a manager at IBM in the ILOG supply chain and optimization group. At Northwestern, he teaches a program on operations management and managerial analytics in the McCormick School of Engineering's Masters in Engineering Management (MEM). He teaches optimization in Northwestern's Master of Science in Analytics program. He holds an M.S. and Ph.D. from Northwestern University in Industrial Engineering and Management Sciences.

Derek Nelson is currently a senior principal at OPS Rules and an Adjunct Professor at Northwestern University. At OPS Rules, Derek leverages analytics to help companies improve operational performance. Prior to OPS Rules, Derek held consulting, product management, and technical sales roles in optimization and supply chain software for LogicTools, ILOG, and IBM. At Northwestern, Derek has taught service operations management to undergraduates in the Industrial Engineering and Management Sciences department and will soon be teaching in the Master in Engineering Management (MEM) program. Derek holds an M.S. in Operations Research from Cornell University.

Preface

Upon first hearing the term *analytics* used in common vernacular, we the authors were unimpressed. After all, we had spent our careers doing analytics. Our undergraduate and graduate degrees focused on topics like mathematical optimization, probability, and statistics. In our careers, we analyzed data all day long: creating databases and reporting systems, working with a variety of tools (including spreadsheets, databases, reporting tools, statistical tools, optimization tools, and specialized engineering tools) to test, manipulate, and understand data in meaningful ways that would allow our company and others we worked with to make better decisions about their business.

We never labeled what we did as "analytics," but surely if anyone was doing analytics, we were.

The term's popularity continued to rise. Soon, it seemed like the term had caught fire. People were using it frequently, it was popping up in marketing campaigns, and whole companies were relabeling themselves as "analytics" specialists. Admittedly, we fell into a common trap of initially resisting the term, basically taking the position that we were seeing yet another example of "Everything old is new again." Don't be fooled, we thought: Analytics is simply what we old-timers call math.

Soon, however, it became clear that this was not just a new term but a movement. And it quickly overwhelmed our ability to resist it. We would see a software vendor say that if you implemented that company's solution, you would be doing analytics. We read about firms in Silicon Valley saying that every firm had to be doing analytics or they would not succeed.

As we read more articles and talked to more people, we realized that the term *analytics* was being used in very different ways by different people. We also saw that software companies and certain types of products seemed to hijack the term and use it as if anything outside of what they were doing wasn't analytics. This didn't sit well with us professionally. Even though we knew that what we were doing was analytics, we feared that people wouldn't recognize it as such because so many were using the term inconsistently.

Eventually, we took a step back, got humble, and admitted to ourselves that perhaps there was more to the story than just the topics in which we had deep experience. There *was* something new about analytics. But before we started researching this topic, we couldn't place our finger on what exactly that was.

In other words, there was not a good definition and description of the field of analytics. If someone wanted to know what this analytics movement was about, it was hard to find a complete answer. To make matters more confusing, the term *Big Data* quickly came on the scene and seemed to be used as a synonym for *analytics*.

We figured that if we were confused, others would be as well. And if managers were now being asked to do more with analytics and Big Data, they needed to truly understand what that was. So we decided to write this book to help managers and analysts everywhere better understand the analytics movement.

We wrote this book for people interested in learning what the analytics movement is all about. We wrote it for people who know that their organization needs analytics to improve but needs to cut through the hype, find a clear definition, and better implement analytics solutions. We wrote it for consultants and software providers so they can deliver better results for their clients. We wrote it for people who are tasked with evaluating analytics solutions but may not know where to start asking questions. We want to see more people, in all kinds of organizations, doing better analytics. We wrote the book from the viewpoint of what a manager needs to know about analytics. Since analytics is technical, we do go into some technical details. But we do so only to help managers gain intuition and insight. And the field is much too large to cover everything in a book like this. We had to pick and choose which topics to cover and to what depth. Certainly, this list could have been different. But, hopefully, we have provided enough information and references that you will be in a better position to explore topics that are of more interest to you.

We hope that a lot of different people get value from this book. We hope managers will be in a better position to evaluate and run projects. We hope that specialists in certain areas will better see how their specialty fits in with other specialties. And we hope that this book will inspire or give you some good ideas to make a significant difference in your organization. Done well, analytics can do everything from help companies create new strategies and save money to help healthcare organizations save lives.

This page intentionally left blank

1

What Is Managerial Analytics?

Confusion About the Meaning of Analytics

The field of analytics (and its frequently used synonym *Big Data*) has captured the imagination of managers everywhere.

Companies often publicly claim that they are committed to analytics. General Electric recently announced a big push into "analytics" to take advantage of the data generated by its own industrial machines.[1] IBM and Ohio State University announced the creation of an analytics center that is projected to employ up to 500 people doing analytics work. Google touts its own internal analytics for its superior searching capabilities and offers its services to others to track their own websites. This list of companies and their newly publicized analytics capability could go on and on.

Not wanting to miss the excitement, consulting and software companies are also touting their analytics services and products. The more skeptical may think that some of these firms didn't change anything except their marketing messages to include the word "analytics." But even the skeptical view does not refute the fact that there is clearly a lot of demand for analytics—otherwise these firms would not be trying to jump on the bandwagon in the first place.

The term *analytics* has even made it into the popular press. It is common to see it used in the business section and even on the front page of the newspaper. The movie *Moneyball* (based on the book of the same name)—showcasing how the clever use of statistics helped the Oakland A's create a low-cost winning team—might be most responsible for the term *analytics* catching the general public's eye.

Those following politics closely have also likely read about President Obama's re-election campaign's heavy use of analytics to target the voters most likely to vote for him. These forays into mainstream culture (sports and politics) furthered the public's interest in analytics.

Responding to the demands of employers and students' desires to be employable, universities have started to offer degrees in analytics. It is very likely that you have come across articles in major publications that talk about the demand for people with analytics skills.[2]

But just what *is* analytics? Companies and journalists are better at using the word *analytics* than at telling you what it is. Without a definition, how do you know if what you are doing is analytics? How do you apply analytics to your situation? How do you know if your vendors are really even selling you an analytics solution?

Whenever a word has such good connotations yet is poorly defined, it is at risk of becoming just another buzzword that will be forgotten when the next one comes around. The CEO wants more analytics projects, so managers put the word *analytics* into all their project titles, without substantively changing anything they are doing. Vendors realize that they can rename what they've always done as *analytics*, and it will sell better. And if a certain group of products become associated with the word *analytics*, then the companies selling those products have no incentive to clarify; why clarify the term if it is working to your advantage? Part of the appeal may very well lie in the mystery of analytics to some people: It sounds complex and promises great returns, so we better do it!

If you are serious about applying analytics correctly, how do you know where to start? How do you know when and where you are going to get value from analytics? Will any analytics projects do, or are some better than others?

Adding to the confusion, the term *Big Data* is starting to be used alongside *analytics*. The term *Big Data* is also not well defined. For example, how is Big Data different from a large data set? And does analytics only apply to Big Data? Do *analytics* and *Big Data* mean the same thing?

This book cuts through the confusion and gives you clear definitions of *analytics* and *Big Data*. You'll get a look at this newly defined field so you will have a deep understanding of the term.

We take the position that analytics is more than just a buzzword or a fashionable trend. Analytics, performed well by capable people, can bring tremendous value to your company or organization. And it applies to companies and organizations of all shapes and sizes. It applies almost everywhere—to large and small companies, non-profits, and educational institutions. It applies to healthcare and medicine, government agencies, science, law enforcement, and the military. Analytics projects can be started by the head of an organization, a manager of a department, or even a single individual. But analytics can bring value only if you and the people in your organization know what it is, can communicate it clearly, and apply it correctly.

What Is Analytics?

The term *analytics* has been used for a long time. So what has made its use so popular now?

The internet companies of Silicon Valley have helped. They use the term *analytics* to refer to keeping track of who is clicking on your website, which pages they visit, what they buy, and so on. But it is hard to imagine that the term has spread like it has if it's just about tracking website performance.

Some companies selling reporting systems (or business intelligence software, if you want to use the industry term) have also helped. These companies claim that analytics is the ability to report on your data in easier and more powerful ways. But again, the term wouldn't be so popular if it were just about reporting.

We believe that the term entered the mainstream business vernacular when *Harvard Business Review* published the article, "Competing on Analytics," by Thomas Davenport in January 2006 (and then a book by the same title).[3] In this article, Davenport highlights how companies like Amazon, Marriott, Harrah's, and Capital One "have dominated their fields by deploying industrial-strength analytics across a wide variety of activities."

This was a tipping point for the term *analytics*. The article really shows that you can apply analytics to a wide range of problems. And it shows that it isn't just a niche area (like tracking website visitors or creating better reports). It's bigger than that.

But we still haven't actually defined what exactly analytics is. The Davenport article gives examples of firms solving specific problems. For example, Marriott uses analytics to set the optimal price for rooms, and Capital One uses it to analyze experiments with different prices, promotions, and bundled services to attract the right customers. But the article only defines *analytics* as the ability to "collect, analyze, and act on data." In other words, according to this article, analytics is about using data to make better decisions. This definition, while correct and compact, does not give much guidance. Haven't managers always talked about using data to make decisions? They have, but there is now much more data available. Simply using more data to make more decisions doesn't really help us define *analytics*.

The contribution of the Davenport article isn't that it defined analytics. Rather, the article helped create the analytics movement. That is, it introduced the idea that a lot of people are solving a lot of different problems using a lot of different tools—and that all these tools are being called "analytics."

The article sparked the idea of analytics as a field (like the field of computer science or of chemistry). As proof of this, many universities have started to offer degrees in analytics. Professional organizations dedicated to data analysis, like INFORMS, have also started to shape the field.[4]

What has emerged from academic and serious business thinkers is a definition of analytics that categorizes the different objectives when using data to make better decisions.[5] We will stick to this emerging definition of analytics throughout this book. **So, here is the definition of *analytics*:**

> Analytics is the collection of disciplines that use data to gain insight and help make better decisions. It is composed of ***descriptive analytics*** to help describe, report on, and visualize the data; ***predictive analytics*** to help anticipate trends and identify relationships in the data; and ***prescriptive analytics*** to help guide the best decisions with a course of action given the data you have and the trends you expect.

Another way of looking at this is to say descriptive analytics aims to provide an understanding of what happened or is happening, predictive analytics aims to tell you what will or may happen next, and prescriptive analytics aims to tell you what you should do. Each of these areas of analytics can be broken down further—which we do later in the book—and different tools and techniques are applied to each.

This definition of *analytics* will hold up over time. It is specific enough to give meaning to the term, while broad enough to allow for future development. New subcategories are likely to come into existence, new tools will surely be developed, and people will gain new insights. This definition gives you a way to understand those changes and how they fit in.

This book dives into each of these areas of analytics to give you more insight. Keep in mind that although we may cover a certain subcategory of analytics or a single tool in just a couple paragraphs, there could be entire university departments, professional organizations, and companies dedicated to just that subcategory or tool. We are by no means understating the importance of that topic. Rather, our goal is to provide you with enough insight so that you can better understand the field of analytics.

To help understand the types of analytics further, let's explore a few examples.

Examples of Descriptive Analytics

A great early example of descriptive analytics comes from John Snow's work during London's 1854 deadly cholera outbreak. The data on where people were dying was readily available but wasn't helping anyone contain the outbreak. But Snow decided to plot the deaths on a map to see if he could get additional insight (see Figure 1.1). In this map, the small black dots represent the residences where people died from the disease. When Snow and others who lived in London at the time looked at the map, they could see clearly that the deaths were centered around a certain water pump.[6] (As noted in the "Endnotes" section, a modern version of this map, creating with modern mapping tools, shows the data even better.) Looking at the data this way helped narrow the search for the source of the cholera outbreak to a particular water pump.

Of course, in reality, it took quite a long time to convince the skeptics that the water pump was the source. But, in the end, Snow was correct, and his map played an important part in convincing people. Some even say that this was the start of the field of epidemiology (the science of studying the patterns and causes of diseases). In this case, Snow visualized data in a new way, which led to a much greater understanding and helped convince others. This is the power of descriptive analytics.

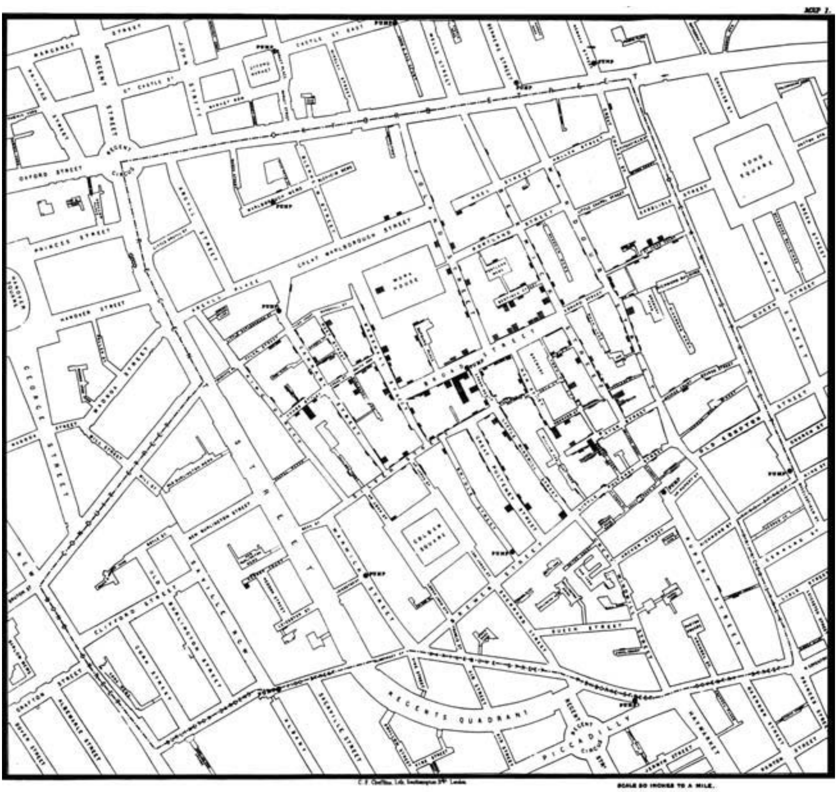

Figure 1.1 John Snow's 1854 Cholera Map

Interestingly, Snow's idea is being applied in a modern way in Lahore, Pakistan, to help prevent the spread of the mosquito-borne disease dengue fever. The city health team is using a smartphone app to record the presence of infected mosquito larvae and plotting this information on a map along with known infections. This information describes the extent or potential for dengue fever, and officials can use it to help determine where to spray. It is a clever update of Snow's idea and shows the power of geographic visualization.

The case of Circle of Moms[7] shows that a little descriptive analytics can lead to big strategic decisions. There was a Facebook application originally called Circle of Friends that had gone viral shortly after it was created and had grown to 10 million people signed up by 2008. Although it looked like they had a potentially popular social website, the founder and CEO realized they had a problem. Almost no one was actually using the application. The founder knew he needed engaged users—not just a lot of people signed up—to have a company with real value. Simple descriptive analytics came to the rescue. The founder began to spend time looking at their data on the users they did have. When he investigated the data, he found that one subgroup was using the application much more than others: moms. The moms were more engaged, had longer conversations, posted pictures, and did other things that indicated they were a happy community group. So the Circle of Friends management team decided to modify the entire business and make it Circle of Moms. Having the ability to look at data and understand it in new ways can lead to big strategic decisions like this one.

Examples of Predictive Analytics

You often interact with predictive analytics through consumer websites, although you may not have known it. Netflix and Amazon are now widely known for their predictive analytics. Based on your ratings (or your purchase and viewing history) and the ratings of others, Netflix and Amazon recommend additional merchandise that you will probably like. It is a bit difficult for an outsider to know how valuable a good prediction system is to these companies. However, we know that Netflix must believe it is pretty important, based on the fact that they conducted an entire contest (available to anyone except Netflix employees or close associates) solely designed to create an improved recommendation algorithm. In 2009, the winning team, which was able to demonstrate a 10% better recommendation system, was awarded the $1 million prize.

Many websites are running tests to predict what designs will work better. In 2008, for example, the Obama campaign was designing its website so that it could maximize the number of people who signed up to be on its email list. Once they were on the email list, people could be contacted for donations or to volunteer. The design team had a choice of several different buttons with different labels (like "Join Us Now," "Learn More," and "Sign Up") and about six different pictures or videos. Instead of arguing about which combination was best, they ran a test. They were getting enough hits to their website that they could randomly assign a visitor one of the different combinations and then track what happened. They ran these tests against a control (the existing design) to predict whether the change would have an impact. In the end, they found a combination that they claimed led to a 40% increase in the number of people who signed up to be on the email list. They claimed that this led to a big increase in donations and volunteer hours.[8]

Examples of Prescriptive Analytics

To borrow an example from the book *The Optimization Edge*,[9] you interact with prescriptive analytics when you use a GPS system or an online map to get directions. You enter the start and end points, and then the program tells you (or *prescribes*) how to get there. Most people don't know that mathematical optimization algorithms are what allow this to happen.

Although you might not think about it, matching kidney donors and recipients is also a prescriptive analytics problem.[10] A person needs only one healthy kidney. Therefore, a healthy donor can donate one healthy kidney to someone who has no functioning kidneys. The alternatives are dialysis and getting a kidney from a deceased person. Waiting for a deceased donor can take a long time, though. Also, the quality and length of life are much better if a recipient receives a kidney from a live donor. Knowing this, a person in need of a kidney can sometimes find a family member or friend who is willing to donate a healthy kidney. However, the problem is that there is a good chance the kidneys won't be compatible; just because the donor and recipient are family members or friends doesn't mean they will be a match. Say that you need a kidney, and your brother is willing to give you one, but you're not a match. If you both get into a database of donor–recipient pairs, you can be matched with a compatible donor, and your brother can be matched with a compatible recipient. You don't actually get your brother's kidney, but he gives up one of his kidneys so you can get one from someone else.

To make a kidney donor–recipient database work, matching organizations use mathematical optimization techniques to prescribe matches. That is, they look at all the possible combinations of matches and pick the best ones that match up the most people and provide the highest possible compatibility matches. The mathematical optimization engine makes this possible. Without it, looking at all the combinations would be impossible.[11] Optimization helps connect many more people and improve many more lives than otherwise would be.

The following examples show companies using all three types of analytics—descriptive, predictive, and prescriptive—to improve their efficiency. As you'll see with both the DC Water and Coca-Cola cases, a lot of the value of analytics results from combining different types of analytics together to come up with a solution that just wasn't possible several years ago.

An Example Using Descriptive, Predictive, and Prescriptive Analytics: DC Water[12]

DC Water, the water company for the Washington, DC, area, serves more than 2 million customers with several thousand miles of pipes and maintains nearly 10,000 fire hydrants. The average age of the pipes is over 75 years. DC Water documented its efforts with IBM to go from a company that used mostly paper records and limited data to an organization that improved performance with descriptive, predictive, and prescriptive analytics. This case highlights how each subcategory of analytics offers value and that they can all complement one another.

By using descriptive analytics, DC Water mapped the location of all the city's fire hydrants. Just visualizing the hydrants allowed the company to create better maintenance plans; before it did, it had been difficult to make sure every hydrant was being properly maintained. DC Water also added extra sensors to water pipes to better monitor water usage and look for anomalies.

With so many aging pipes, failures were a big problem. In the past, DC Water simply reacted to the failures. With predictive analytics, the company could now look at factors such as the age of the pipe, soil conditions (gathered from better descriptive analytics), pressure on the pipes, nearby problems, and other factors and use statistical models to predict when pipes would fail. This allowed DC Water to address potential issues before they became actual problems. It also helped the company prioritize preventive maintenance.

By using prescriptive analytics, DC Water could better route maintenance crews to fix trouble tickets, which increased the productivity of the maintenance team while driving down fuel cost. By knowing the location of each hydrant, existing work order, or preventive maintenance project, DC Water could better route the trucks and crews to the best locations.

An Example Using Descriptive, Predictive, and Prescriptive Analytics: Coca-Cola Orange Juice Plant[13]

Businessweek published an article on Coca-Cola's new state-of-the-art orange juice plant. The goal of the plant was to produce high-quality and consistent orange juice year-round. The article showed how Coca-Cola was using analytics to accomplish this.

As an example of descriptive analytics, Coca-Cola used satellite images of the orange groves in Brazil to determine when different fields were ready for harvest. By being able to see the orange groves in a new way, Coca-Cola was able to better harvest the oranges for increased quality.

Coca-Cola used predictive analytics to predict the quality of oranges coming from different fields by looking at different weather patterns. This helped determine what types of oranges would arrive and whether the company would have to acquire oranges from other locations (if a particular location was likely to have a low-quality harvest). Coca-Cola also analyzed the chemical components of oranges to predict the taste and the range where people could detect a difference. For example, to make up numbers to illustrate a point, Coca Cola would need to determine whether the sweetness range needed to be between 100 and 200 or between 155 and 175 so they could maintain consistency.

Finally, Coca-Cola used prescriptive analytics to determine how to blend the oranges of different quality and characteristics to get the desired output. The mathematical optimization looks at all the potential oranges available and comes up with the mix that gets exactly the right combination of all the different chemical characteristics at the lowest cost.

So, What's New?

Based on our definition of *analytics*, you may be thinking that many of the disciplines within analytics are not new. You are correct. Many of them have been around for years. But many things are new. We'll cover these in more detail as we move through the book. The following list will give you a flavor for where we're going.

First, there is a now a realization that all these disciplines are related and part of the larger theme of analytics. New solutions become possible as you combine different types of analytics to solve a problem.

Second, the proliferation of data has opened the possibility of asking many new questions and uncovering many new insights and trends. Analyzing and summarizing that same data, however, requires sophisticated tools and methodologies because of the data volume and complexity. At the same time, to take full advantage of the opportunity presented, organizations need to be able to deploy tools and methods more widely; in the past, these tools and methods were stuck in the corner of an organization, run by experts. The more people who can look at and analyze the data, the better your chances of finding interesting trends and running your business better. In addition, this proliferation of data has unleashed a new wave of creativity. Early on, website developers realized that they could combine two or more different services or products to create a new one. This was called a "mash-up." More abundant data has made mash-ups accessible to more managers. They can pull data from many different sources to solve new problems or to solve an old problem in a new way. Creating mash-ups requires creativity.

Third, the proliferation of data has given people incentive to create new tools, revisit old tools that didn't work with limited data, and apply old tools in new ways. For example, when you have access to the whole universe of data rather than just a sample of it, you can analyze it with algorithms that might not have worked well with the smaller sample set of data. (We'll give some nice examples of this in Chapter 2, "What is Driving the Analytics Movement?") Also, with large data sets, mathematicians are rediscovering old fields. For example, the field of topology has been around as a purely theoretical field for 250 years. Now topology is being used to help people visualize large data sets. Finally, new tools (or updated versions of them), like machine learning algorithms (which we will cover later), are moving out of research labs and into the hands of businesspeople.

Fourth, with a lot of business moving online combined with the fast feedback of social media, analytics makes running tests much easier. For example, a company can show different versions of its website to randomly selected visitors and easily test which version leads to the desired results—such as more sales, more signups, or longer time on the site. With more business being done online, analytics can help make and influence more and more business decisions.

Another way to look at what is new in the field of analytics is to consider how it can change what managers need to do. The former president of a successful online financial services firm summed it up nicely. He said that his job wasn't to figure out what decisions to make. Instead, it was to figure out how the decisions should be made. Once he was confident in how decisions should be made, algorithms could be programmed to make those decisions. The algorithms ran the business. The algorithms determined what webpages to show each visitor, what services to offer the visitor, and what price to charge. Management simply needed to make sure the algorithms stayed up to date and used additional data as it became available.

Finally, and possibly most importantly, the attitude toward analytics is new: With the abundance of data available and the abundance of tools to use that data, managers realize that more and more decisions can be improved through the use of analytics. If they don't take advantage of that, they risk falling behind.

What Is the Best Type of Analytics?

Often, when people explain the three types of analytics, they present a diagram that shows descriptive, predictive, and prescriptive analytics building on each other. That is, you need to do descriptive analytics first, then predictive, and finally prescriptive. The diagram then usually suggests that descriptive analytics is the easiest but adds the least amount of business value. On the other end of the spectrum, prescriptive is labeled as the hardest but yielding the most potential business value. Figure 1.2 shows an example of this type of diagram.

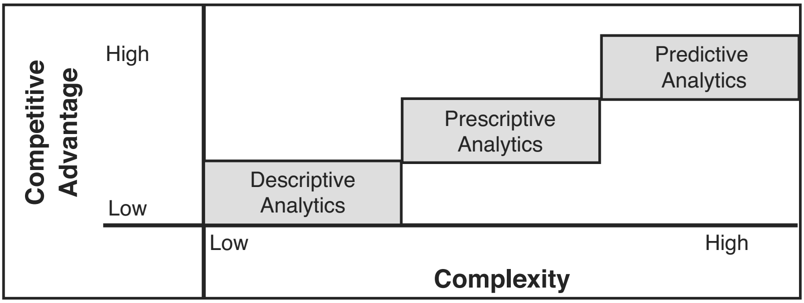

Figure 1.2 Sample of a Typical (but Misleading) Diagram of the Types of Analytics

This diagram is visually interesting and can provoke a good discussion. And we, the authors of this book, used such a diagram extensively in the past. However, we now think that this diagram is misleading. In some cases, the diagram can be true. But, it is not universally true. And it may be true in only a small number of cases.

The value of the different types of analytics is tied to the problems you are solving, not the techniques themselves. Each type of analytics can have relatively minor impact on a business or completely change the business.

For example, if your descriptive analytics project is to just come up with a better understanding of what type of product is sold in what geographic area, you will likely get some interesting insight, but it is unlikely to change your business. If, on the other hand, your descriptive analytics project uncovers deeper insight about your customer, it could cause you to change your entire business, as in the case of

Circle of Moms. This represents a huge strategic change. As another example, good descriptive analytics applied to data sets concerning cancer patients has uncovered potential life-saving treatments for different types of patients. It would be hard to argue that this type of descriptive analytics isn't extremely valuable.

Predictive and prescriptive analytics work the same way. You could have many nice projects that are relatively small in terms of the overall strategy of the firm (but could still be good for that segment of the business) and you can have projects of strategic importance. For example, a small predictive project may involve improving forecast accuracy by 10% for certain items. Certainly this is nice, but it's unlikely to change the business. On the other hand, Amazon's use of predictive analytics to make buying suggestions changed the nature of retailing.

For prescriptive analytics, we could imagine that the routing of trucks for DC Water saved some money but didn't change the business. On the other hand, Coca-Cola's use of prescriptive optimization to blend orange juice to improve the quality of the product could have significant strategic impact on the product. In another example, Indeval, Mexico's central securities depository, uses prescriptive analytics to help settle securities. The system determines how to match buyers and sellers and clears $250 billion per day, saved $240 million in interest over 18 months, and made the market much more liquid. Indeval says that the system was in place during the stock market crash in 2008, and the extra liquidity in the system allowed people to exchange securities more often during a single day—which was much better than in other countries, like the United States, where the lack of liquidity meant people were stuck holding securities for an extra few hours as the price of the security rapidly decreased. Indeval's use of prescriptive analytics is obviously very strategic.

It is also impossible to say which type of analytics is easiest to implement. You can have a simple descriptive analytics project where you load the data you have into a better reporting tool and immediately gain insight. Or you could spend two years implementing a full-blown descriptive analytics system that gives you access to every bit of data in your organization. Likewise for predictive and prescriptive analytics: You can do good work in an Excel spreadsheet or you can custom build systems that require years of effort and huge teams of people.

Finally, there is no particular order in which to implement these systems. You might think you need to do descriptive analytics to understand what is going on in the business first. But if you have data (which most firms do) and know what problem you want to address, you can skip descriptive analytics and start directly with a predictive or prescriptive project. And, in practice, this is often what happens. The types of analytics projects you do depend on the business issues you need to address, the value of the project, the ease of doing the project, the skills of your team, and many other factors.

We wish there were a simple roadmap that you could follow to get the most out of analytics. But the wide range of analytics applications and tools is what makes the field so fascinating and rich. You have many different options and have to pick the best approach for your business for each project (big or small).

What Is Managerial Analytics?

We use the term *managerial analytics* as the title of this book, so we should define what it means. We define *managerial analytics* as what a manager needs to know about the field of analytics to make better decisions. So we are using the term to refer to understanding the field of analytics from a manager's perspective. Or, in other words, *managerial analytics* helps prevent you from being fooled or confused by the many different analytic terms and solutions offered, helps you see what is possible, and helps you do more on your own.

Managerial analytics is about understanding the difference between descriptive, predictive, and prescriptive analytics. It is about understanding the different tools and where they fit in. It is about cutting through the buzzwords to help you better understand solutions.

Hopefully you are now sensing how large the field of analytics really is. You can get an advanced degree in very specific niches within the analytics field. As a manager, you cannot understand every nuance. You should, though, understand what problems the different disciplines of analytics can solve. And you should know what types of decisions a particular project will help you with.

Another way to think about this is that you will often be presented a single tool as a way to do analytics. We want this book to help you, as a manager, cut through the confusion and determine which type of analytics you need and for which types of projects a tool may provide value. No one solution will cover the full range of the field of analytics. This knowledge should help you select analytics projects that meet your objectives and then guide successful implementations.

Managerial analytics is not devoid of technical material. As the importance of data and analytics increases, you will have to be more comfortable with the technical aspects if you want to succeed as a manager. While you won't be expected to understand all the nuances of different disciplines, you should understand the limitations of a given analytics solution—and what that solution won't do. Many times, an analytics solution will be presented by someone pushing an idea, and it will be impossible to tell what the solution won't do. Vendors commonly write descriptions of their analytics solutions that make it seem like a single solution will solve all your problems. If only it were this easy.

This book covers enough of the key technologies behind descriptive, predictive, and prescriptive analytics so that you will know if a project is on track, you will know what experts you may need, and you can better understand the details of the solutions being presented to you and have a healthy understanding of their limitations. In other words, with this understanding, you are less likely to be fooled or confused by vendors and co-workers using the term *analytics* in a vague way.

But managerial analytics is about more than not getting fooled. It is also about the art of the possible. The analytics movement is real because it produces real value. To capture this value, however, you need to know what is possible. So managerial analytics is about understanding how analytics can apply to your business. One of our goals with this book is to get you thinking about your company or organization in a new way. That is, by seeing different examples and understanding the different areas of analytics, you may uncover opportunities for adding value in ways you hadn't thought of before—and these new ideas may turn out to be much more important than current projects you are working on. If you want your entire company or just your small department to make better use of analytics, the more you and your colleagues understand what is possible, the more value you will find.

Finally, not everyone has access to a large IT department and teams of analytics experts. In fact, most people don't. So don't think of managerial analytics as just something that big firms do. Instead, managerial analytics is for organizations of all sizes. Managerial analytics is about giving you enough information that you can get started with what you have. From what we have seen, most managers have plenty of room to use the data and tools they have in order to make better decisions. Don't wait for big projects; you can start now.

The rest of this book is devoted to helping you understand managerial analytics—analytics from a manager's perspective. This will help you whether you are doing a single project within an organization or rolling out analytics solutions to your entire organization.

It will help you understand the definition, help show you how the types of analytics work together, and present you with practical applications that are applicable to a wide range of organizations. It provides cases and examples to solidify the ideas throughout. And, since analytics is a technical field, this book also discusses the technical aspects of analytics. You can skip these sections without losing the flow of the book. But, as a manager, the more you understand the technical aspects, the better you will be able to apply analytics.

Is Competing on Analytics a Strategy?

Davenport's article, "Competing on Analytics," was not written to define analytics. It was written to show that firms that are dedicated to using analytics to tackle the biggest issues and using analytics throughout the firm could do significantly better than other firms in their industry. In the article, he also discusses ways a firm can structure itself to be an analytics competitor. It becomes clear from the article that analytics could be a highly strategic tool for any firm.[14]

An article in *Analytics Magazine* stated, "In many ways, business analytics is the next competitive breakthrough following business automation but with the goal of making better business decisions, rather than simply automating standardized processes."[15]

And you can find many other similar claims.

Based on all this, it might be easy to declare that competing on analytics can be classified as a strategy. A CEO may try to claim "We will compete by being the best at analytics."

Ben Reizenstein, from Northwestern University's Kellogg Business School, actually posed this exact question: Is analytics a strategy? Or, is it just a tool to help you do what you do, but better? He was leaning toward the latter. That is, with all the hype around analytics (much of it deserved), we might have lost track of the idea that analytics can help support your strategy. So, maybe it is not a strategy but a great tool to help you execute your strategy.

For example, "Competing on Analytics" uses Marriott as an example. For a hotel chain, the strategy might be to provide the best business hotels or the best resort hotels, or to be in every market or to be only in large urban business districts. The use of analytics is a way to help the hotel chain really execute its strategy, but its commitment to analytics is not necessarily the strategy itself.

If analytics is a strategy, then you have to wonder if every analytics project has value. If analytics supports your strategy, you can then judge analytics projects as they relate to executing your strategy.

In Davenport and Harris's book, *Competing on Analytics* (the follow-up to the article), they mention that analytics is not a strategy but that it supports developing certain business capabilities.[16]

On the other hand, consider again the hotel example. You could imagine a scenario where a hotel decides to be the best business hotel and will use analytics to do just that. This could be contrasted with a strategy of being the best business hotel by focusing solely on hiring service-oriented people or having a unique sales force to win corporate clients, or to focus on buying distressed hotels in good markets to keep real estate costs low. The list could go on. There is nothing that dictates that a firm needs to use analytics to be better. In this sense, analytics is an integral part of the strategy.

It dictates the type of people the firm hires and the investments the form makes in IT, for example.

Davenport and Harris mention that the companies that are best at analytics mention it in their annual reports and press releases, and it is highly visible to the top of the organization. They even mention that the firm's strategies are built around analytics—suggesting something possibly a bit closer to a strategy than just a supporting role.[17]

Another way to look at this issue is to divide strategy into the business strategy (what the business is trying to be) versus operational strategy (how to achieve the business strategy). With this view, clearly analytics is an operational strategy that supports the business strategy.

We want to present you with a balanced view that neither over- nor undersells analytics. Our aim with this discussion on strategy is not to provide an answer but hopefully to stir a lively debate in your organization about the role that analytics will play.

This page intentionally left blank

Index

Numbers

A

D

N

O

R

S

T

Y

Z