CHAPTER 1

Introduction and Background

THE POWER OF INFORMATION

Revolutions come and revolutions go; while their influence may be obvious in retrospect, it is rare that we appreciate their impact at the time. The one thing they have in common is disruption at all levels of society. First, the agricultural revolution; then the industrial revolution; and now, 60 years into the latter, we have the information revolution.

Information has always been power, but the past few decades have seen a subtle shift occur, fundamentally altering the way we perceive it. It has been only relatively recently that the amount of data available to us has outstripped our ability to investigate that data. At one point in time, it was arguably possible to have read every written word ever set on parchment or papyrus. While the true number will never be known, it is said that King Ptolemy II Philadelphus set the Library of Alexandria a target of 500,000 scrolls. At the time, this represented the largest collection of knowledge in the known world.

Things change. At the time of this book's publication, the United States Library of Congress had over 33 million books (including other printed material) and 63 million manuscripts. The Internet Archive,

capturing only a subset of all the information contained on the Web, has already cataloged almost 2 petabytes of text and is growing at approximately 20 terabytes a month, in itself a larger amount of information than that held by the Library of Congress.

A little over 2,000 years have passed between King Philadelphus and now. Things are accelerating at such a rate that in another 20 years our staggering statistics will probably be considered equally quaint.

This tsunami of information is a real challenge at every level in society. At a personal level, we struggle to keep on top of everything that is happening around us. Alvin Toffler coined the term "future shock" as early as 1970 to describe the overwhelming and disorienting impact of information overload.[1] And, at a professional level, where we once struggled with a paucity of information we now struggle to pick which pieces of information are important out of the millions of measures at our fingertips. Regardless of where you start, this ever-increasing amount of information has changed the way we view the world, the way we live, and the way we do business.

Dealing with this data deluge requires being smarter. It requires developing the ability to selectively process information based on value, not sequence. It requires, more than anything, the realization that brute force and manual effort are, in the long run, an impossible solution. Quite simply, it requires the effective application of analytics.

Getting to this point has not happened overnight—it has taken decades of research and thought. Among others, Claude Shannon led the charge when he published his seminal work, "A Mathematical Theory of Communication": His proof that information could be quantified and measured was both innovative and revolutionary.[2] Without his research, our world would be vastly different. Among other things, it is unlikely that Voyager would have launched or that the Internet would exist.

Treating information as a measurable quantity has changed the world. While statistics had always been concerned with extracting useful knowledge from raw data, information theory helped encourage the perception that information was more than just insight.

Instead, it could legitimately be seen as an asset in its own right, something of real (and comparable) value.

This intersection of statistically based insight and the realization that information can be an asset has had and will continue to have serious reverberations in the business world. Being smarter has always meant being successful; as far back as the nineteenth century, analytics was already generating competitive advantage.

William Gosset, employed by Guinness, had been struggling to identify the varieties of barley that had the best productive yield. Like medical researchers and biometricians, he was forced to deal with extremely small samples, often as little as 30 or fewer observations. Through a combination of rigorous research and trial and error, he identified a new distribution to help model likely population means, something that gave Guinness a significant competitive advantage through optimizing production efficiencies. By using his distribution to better predict crop yields, Guinness was able to gain a cost advantage over its competitors.

An interesting side note is that despite Guinness having prohibited employees from publishing their discoveries (given their understanding of the importance of analytics as a competitive advantage), Gosset published his work anyway. However, in the interests of keeping his job, he published it under a pseudonym, giving us the well-known "Student's *t*-distribution" that we work with today.

MODERN-DAY MAGICIANS

Today, we face the opposite of Gosset's challenge. While we still sometimes have to deal with fewer observations than we would like, the growing challenge is working out how to deal with the massive amounts of information we do have. As Moore's law suggests, we have seen the number of transistors we can fit inexpensively on an integrated circuit double roughly every two years. Although it is not an exact relationship, we have seen roughly exponential growth in processing power since the early 1950s.

Importantly, processing power is not the only area that has grown by leaps and bounds. Kryder's law suggests that disk storage density

doubles annually, a pattern that has largely held true since the mid-1990s, when it was first suggested. And Butters's law suggests that the amount of data carried by a single optical fiber doubles every nine months.

Combined, these create our future. We have the storage capacity to track ever-increasing amounts of information, often referred to as "big data." We have the bandwidth to support the transfer of this information as needed. And, we have the processing power to extract insight from this data.

Because of this, we are drowning in data. For the first time ever, we have more data than we have storage capacity. In 2008, International Data Corporation (IDC) estimated that the amount of data being generated exceeded our total aggregate storage. Organizations such as eBay have repositories in the petabytes, and there are no signs that this accelerated rate of data retention is going to slow.

The answer does not lie in better processes. Enterprise resource planning (ERP) systems play a critical role in standardizing operational processes, but they do not create competitive advantage or insight. Their focus is instead on establishing process efficiencies. There are obvious advantages to this, one of the biggest being that it reduces cost in the long run. However, because they rely on standardizing processes between organizations, at best they simply set a minimum benchmark. *It is impossible to differentiate yourself if you use exactly the same processes as everyone else.*

Instead, competitive advantage comes from capitalizing on what makes you unique. Every organization is different, and every organization has the potential to exploit that exact uniqueness in a way that no one else can match. Doing this means taking advantage of their single biggest resource: their data.

The people who know how to manage this data deluge are our future. Hal Varian, Google's chief economist, summed it up best: "I keep saying that the sexy job in the next ten years will be statisticians."[3] Being able to translate massive amounts of data into real insight is beyond magic—it is competitive advantage distilled. Nothing else offers an equivalent level of agility, productivity improvement, or renewable value. Being "smarter" than your competitors is not just hyperbole; it is a real description of how significant the impact of

applied analytics can be. If yours is an organization like Zappos.com, trying to understand what over a million people are telling you every day through social media, it is simply impossible to do business without leveraging advanced analytics for text mining and sentiment analysis.

Today, we stand on the cusp of transformation. The organizations that successfully extract the maximum amount of useful information from their ever-increasing data repositories and act on their insights are on the path to success. Those that do not are inevitably set on a track toward mediocrity at best, and abject failure at worst. Much as Henry Ford's assembly line forced handmade production into a niche, the organizations that successfully apply business analytics will increasingly reduce the relevancy of those that do not.

However, success requires more than just knowledge of statistics or ways of dealing with big data. Execution is essential, but without a plan and commitment, little happens. It also requires an understanding of how analytics translates to competitive advantage. It requires an ability to enact change within an organization. And, it requires managers to sell the value of analytics.

While analytics has a long and hallowed history of creating organizational value, it remains a relatively opaque discipline. For many organizations, the insights created by advanced analytics are often tied to individuals, not processes, making it hard for them to extend the value to other areas of the business. While there is rarely any disagreement that any organization should be smarter in its decision making, the linkage between numerical analysis and value creation is rarely understood. When it comes to investing in building these capabilities, many organizations are reluctant to take the leap of faith that is apparently necessary.

It does not need to be that way.

THE SECRET OF SUCCESS

When one looks globally, some organizations seem not to mistrust analytics. Somehow, they seem to create renewable value through applying their competencies across many different business problems. Somehow, they succeed, not once, but repeatedly. And somehow,

they translate their experience and skills into sustainable competitive advantage. Whether one is looking at Marriott, Canada Post, Sainsbury's, or Telstra, some organizations consistently manage to deliver significant returns through their application of business analytics.

Possibly the best-guarded secret in business analytics is that in practice, their success comes down not only to organizational culture but also to the ability of their managers to successfully sell the value of analytics. As researchers such as Thomas Davenport and Jeanne Harris have rightly pointed out, overall success can often be linked to a variety of factors, including organizational structure, management commitment, and successful strategic planning. However, it is often "where the rubber hits the road" that the greatest impact can occur.

Change comes from two directions: top down or bottom up. If the organization already has the right management culture, using insight as a competitive advantage is relatively straightforward. As Jack Welch is reputed to have said, "An organization's ability to learn, and translate that learning into action rapidly, is the ultimate competitive advantage."

In these organizations, applying business analytics is relatively easy. There is management commitment to the use of insight, there is an understanding of the role analytics plays in creating competitive advantage, and there is often a culture of continuous improvement. For those of us lucky enough to work in this context, life is easy.

Unfortunately, most of us do not work with these types of organizations. Instead, we work in environments that are unfamiliar with the use of analytics. Environments that struggle to differentiate between intuition and fact, often taking pride in making decisions based on "gut feel." Environments that, as frustrating as it can be, cannot understand the value that business analytics provides.

This book is written to help those who want to change the environment in which they operate.

Analytics is a multidisciplinary activity: The value from insight comes not from the activity but from the execution. Often, this crosses a variety of departments within an organization. Few analytics groups have responsibility for both the insight creation and the execution of

that insight. Because of this, selling the value of analytics is not just an aspirational goal for managers; it is a necessary criterion for success.

For many managers, this can be challenging. Despite broad interest in business analytics as a discipline, there exist few useful sources that help managers with the practicality of getting organizational commitment to using business analytics. This book aims to fill that knowledge gap.

NOTES

1. A. Toffler, *Future Shock* (New York: Random House, 1970).
2. C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal* 27 (July–October 1948): 379–423.
3. J. Manyika, "Hal Varian on How the Web Challenges Managers," *McKinsey Quarterly* (January 2009), www.mckinseyquarterly.com/Hal_Varian_on_how_the_Web_challenges_managers_2286.